Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 35, NUMBER 4 APRIL, 1945

Recent Studies on the Fluorescence of Glass*

N. J. KREIDL
Bausch and Lomb Optical Company, Rochester, New York

1. THE USE OF FLUORESCENT GLASSES

THE most spectacular example of the manifold interests of the glass technologist in the subject of fluorescence is the field of illumination. In general practice the fluorescence of cold light sources is produced by coating the interior surface of the glass tube with crystalline fluorescent substances. Some research, however, has been devoted to the investigation of the fluorescence of the glass proper. Uranium glass was first observed to be fluorescent and only recently other types of glass containing unusual and carefully selected constituents were investigated.

While the possible application of fluorescent glass to illuminating engineering is still under consideration the phenomenon itself has been developed into a useful tool for the investigation of the constitution of glass. It was conceived that the light emitted by a fluorescent unit in the glass could be utilized as an indicator or "signal" of the location of the fluorescent unit as well as of the nature and the eventual changes of the unit's surroundings.[1]

In optical glass research the technologist developing new optical glasses is interested in the chemical constitution. Insight into the glass structure is procured by such a small number of methods that any new and independent method is welcomed. The use of fluorescence indicators for structural research is based upon the effect of the atomic environment of the fluorescence centers.

The light emitted by a fluorescence center is affected by, and therefore associated with, the intimate structure of its surroundings. Thus, structural changes of the same glass owing to heat treatment or different structures in glasses of different composition affect the energy; that is, the color and intensity of the fluorescence.

Since uv light sources have become readily available they now represent a standard piece of equipment in the research and testing laboratory. Soon they will be standard equipment in the glass plant and in the optical shop. The glass technologist thus will not only examine glasses of unknown character but eventually every piece of glass.

For the selection of the proper source of radiation it is important to realize that the colors and intensities produced depend upon the wave-length of the uv light as well as the character of the glass. Moreover, the glass must absorb the wave-length to which the fluorescent unit responds.

As the more recent sterilizing lamps make the shorter wave-length range easily accessible it becomes apparent that nearly every glass possesses a characteristic fluorescence!

* Presented at the Twenty-Ninth Annual Meeting of the Optical Society of America, New York, New York, October 20 and 21, 1944.

[1] W. A. Weyl, Sprechsaal **70**, 578–82 (1937).

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

2. THE FLUORESCENCE CENTERS

The fluorescence of glass is caused by a relatively small number of chemical units termed fluorescence centers. Light, usually uv light, is absorbed and its energy is used to move an electron to a more distant position with respect to the nucleus. From these unstable levels of higher energy an electron may return, emitting light. In accordance with Stokes' law the energy of this emission is smaller than that of the irradiation. Thus the emitted light has a wave-length greater than that of the incident light. The loss of radiant energy is caused by energy dissipation chiefly in the form of heat and the amount of the energy loss depends upon the nature of the surroundings.

At times it has been believed that the crystalline condition is necessary to create fluorescence. While some fluorescent glasses contain crystalline matter, other glasses entirely free from crystals are equally capable of fluorescence: In fact, almost all glasses fluoresce.

W. A. Weyl classified the fluorescent glasses on the basis of their fluorescence centers,[2] as represented in Table I.

Most of the glasses discussed in this paper pertain to class 4(a), the fluorescence centers of which are cations.

In fluorescent glasses proper (3 and 4, Table I) the fluorescence is affected by structural changes in the glass in the following manner: (1) The intensity of fluorescence is decreased and its wave-length shifted towards red the more the fluorescence center is disturbed by the surroundings. The interpretation of these changes can extend the evidence of x-ray analysis of simple glasses to more complex glasses. (2) Increasing symmetry of the environment narrows the fluorescence bands. This aids in evaluating the structural difference of crystals and glasses of equal composition, especially as a function of temperature. (3) Increasing viscosity and less frequent collisions increase the lifetime of the excited state. Fluorescence of longer duration is observed.[3]

TABLE I. Fluorescence centers in glass. (After Weyl's classification of fluorescent glasses.)

Fluorescence center	Example
1. Active crystals	Crystalline zinc silicate activated by Mn^{++} embedded in a low melting glass in which it is insoluble.
2. Activated crystals	The glass contains Mn^{++} which is adsorbed by the crystal.
3. Atoms or molecules	(a) molecules CdS; (b) atoms Ag;
4. Ions	(a) cations Mn^{++}, UO_2^{++}, Ce^{3+}; (b) anions VO_4^{3-}.

[2] W. A. Weyl, Ind. Chem. Eng. **34**, 1035–1041 (1942).

3. THE CONSTITUTION OF GLASS

The discussion of fluorescence as an indicator of glass constitution requires a brief discussion of the present problems in this field.

The reliance upon recipes, those combinations of unexpendable truth and superstition, was followed by attempts in the 19th century to force the fields of glass formation into chemical formulae. It was, however, found impossible to render glass composition by stoichiometric formulae. In chemistry, formulae of constitution were necessitated by the discovery of different properties in compounds of identical composition. In glass technology the subject of constitution was brought up by the reverse problem: Certain glass properties are independent of the composition but simply characteristics of the "vitreous state."

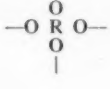
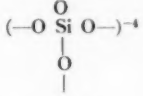
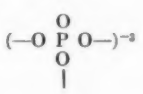
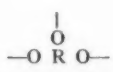
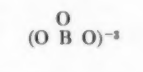
The comparison of glasses with solutions had its merits but the first profound satisfaction was derived from the interpretation of x-ray analysis. On this basis the glassy or vitreous state was explained by the similar geometrical arrangement of various atoms of similar relative charges in space.

This initial approach was fertile. As is customary in crystal chemistry, rules, governing the charge and the apparent size of cations and anions, were set up which fitted well the interrelation of the most frequent one-compound glasses: SiO_2, B_2O_3, P_2O_5. These rules are summarized in Table II.

The structural predictions of Goldschmidt and Zachariasen[4] were proven quantitatively by Warren's[5] absolute x-ray analysis: The structure of a SiO_2 glass, for instance, is that of a network of

[3] W. A. Weyl and E. Thümen, Sprechsaal **67**, 95 (1934).
[4] W. H. Zachariasen, J. Chem. Soc. **54**, 3841 (1932).
[5] E.g., B. E. Warren, J. Am. Cer. Soc. **17**, 249 (1934).

TABLE II. Conditions of glass formation.

Conditions				Examples			
Structural unit	Formula of unit	Radius cation R / Radius anion O	Sum valencies O / Valency R	Formula	Structural unit	Radius R / Radius O	Sum valencies O / Valency R
Tetrahedron		0.22–0.41		SiO_2		0.32	$\frac{4\times O}{Si}=\frac{8}{4}=2$
			About 2 so that each O can be tied also to next R	P_2O_5		0.26	$\frac{4\times O}{P}=\frac{8}{5}=1\frac{3}{5}$
Triangle		0.15–0.22		B_2O_3		0.15	$\frac{3\times O}{B}=\frac{6}{3}=2$

SiO_4^{4-} tetrahedra (four oxygens surrounding one silicon) with each oxygen shared by two silicons.

Oxides following these rules and capable of forming glasses by themselves were termed "glass or network formers." Typical glass formers are the oxides of boron, silicon, and phosphorus.

It is true, however, that almost every element may participate in the glass structure to some extent even though it does not meet these requirements. This was explained by the location of these secondary constituents, the so-called "modifiers," in holes of the network of the "glass formers" proper. Typical modifiers are Na^+, and Ca^{++}.

The modifiers cause a disruption of the strong SiO_4^{-4} network. A certain number of oxygens is no longer shared by two glass forming cations and the loosened structure becomes manifest by properties such as lower melting temperature, lower viscosity, or lower chemical resistance. There is no doubt that the distinction between hole positions and network positions must be conceived as a relative one.

The real meaning of this distinction between glass formers and modifiers is the predominant control of charge distribution in the polyionic glass by the stronger fields of the ions capable of glass formation. The distances of the "glass formers" to oxygen are nearly as regular as in a crystal while the distances of the weaker ions (modifier ions) are subject to greater variation. It is the more regular distances of the glass forming ions which are recorded by x-ray analysis.

It has become evident that neither very strong nor very weak bonds between a cation and its surrounding oxygen ions are conducive to glass formation. (Fig. 1.) Too strong forces attract all coordinated oxygens forming "molecular" and usually volatile compounds. There is not enough residual negative field in the surrounding oxygens to produce continuous linkages. Too weak forces lead to "ionic" and soluble crystalline lattices.

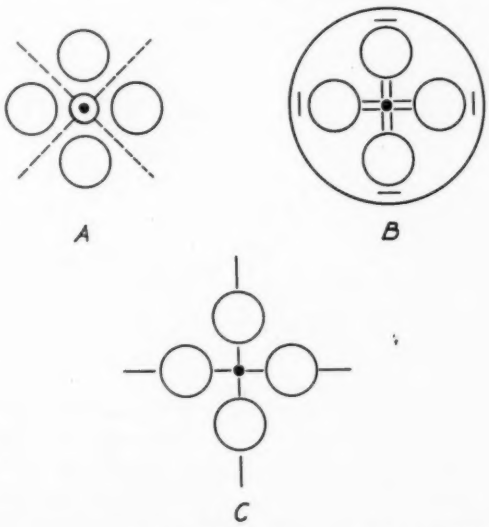


FIG. 1. Simplified scheme illustrating how intermediate forces in binary compounds favor continuous networks of refractory oxides and glasses. *A*. Weak interionic forces: so-called polar compounds, ionic lattices. *B*. Strong interionic forces: so-called molecular compounds; usually volatile at low temperature. *C*. Intermediate interionic forces: tendency to use about half the field strength of the anion to bind the next cation, leading to continuous networks of refractory oxides and glasses.

The rule that the cation should have a charge of 3, 4, or 5 thus being approximately one-half that of the surrounding 3–4 oxygens is a first approximation of the general conditions of field strength leading to a continuous network.

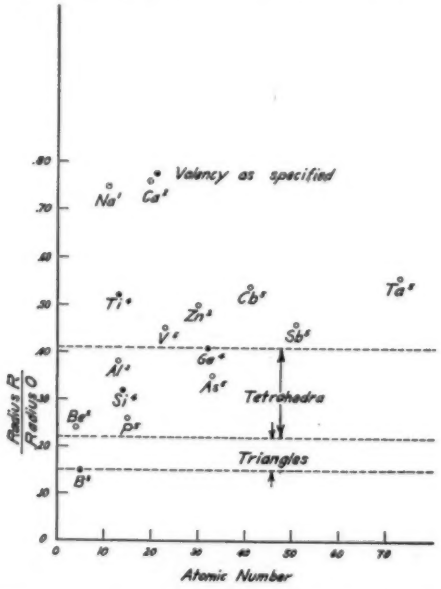


FIG. 2. Apparent radii, calculated from crystallographic data by Pauling and others, of ions capable of perfect or partial glass formation with oxygen. The radii are given in fractions of the apparent radius of oxygen in silica. Dashed lines represent the limits of tetrahedron and triangle formation as specified by Goldschmidt's classical rules. Tetravalent Si^{4+} and Ge^{4+} are within these limits, fulfill the necessary valency requirement and form perfect glasses. Pentavalent P^{5+} also forms a glass with oxygen. Trivalent Al^{3+} and divalent Be^{2+} are capable of participation in glass networks. B^{3+} is located at the boundary of triangle formation and forms a perfect glass B_2O_3. The influence of the actual field instead of the simple classical rule becomes manifest by many ions being capable of participation in glass networks outside of these limits: Ti^{4+}, V^{5+}, Zn^{2+}, Cb^{5+}, Sb^{5+}, Te^{5+}. The modifying ions Na_2 and Ca_2 are inserted for the sake of comparison and contrast.

For a more precise description of the structure the actual field and not the symmetrical charges of hypothetical spheres must be taken into account. The influence of charge distribution, polarization, and polarizability overrules the simple concept derived from the first approximation. It becomes manifest specifically in the case of the ions of non-noble gas character in the subgroups of the periodical system.

It was always suspected that trivalent ions such as Al^{3+} participated in the glass network. But by means of the additional facts just presented it may be understood that even divalent ions like Zn^{2+}, Pb^{2+}, and, at least to a certain extent, Ba^{2+} can assume the role of glass forming cations because of the field created by a particular electron distribution existing in the ion or promoted by polarization.

The arrangement of the apparent radii of ions capable of some participation in the glass structure proper in the vicinity of the radii of ions specified by the initial approach is shown in Fig. 2.

Table III shows some older and newer glasses with constituents other than silica, comprising PbO, Al_2O_3, P_2O_5, and some rare element oxides.

This analysis of constitutional problems shows that complex glasses may comprise the entire periodical system of elements. In these complex glasses the fluorescence of the same center varies considerably.

4. FLUORESCENCE AND CONSTITUTION OF GLASS

Assuming that the main constituents of a simple glass are (1) oxygen, (2) a glass forming strong cation like silicon, and (3) a modifying cation like sodium, the intensity, the wavelength, and the width of the fluorescence band are found to be affected by any change in these constituents.

The intensity is affected by a change in the nature of the glass forming as well as the modifying cation, and by the ratio of the one to the other.

In general it appears that (a) a high percentage of glass forming cations, (b) strong glass forming cations, and (c) modifying ions of low field strength favor intensive fluorescence.

The most suitable fluorescence center for intensity studies is the uranyl group. Uranium

TABLE III. Some glasses containing glass formers other than silica.

Name	Glass former	Main property
Pyrex	Si, B	Low expansion, high resistance
Flint	Si, Pb	High refraction, high dispersion
Alumophosphates	Al, P	Insulation, low melting
High index glass according to Morrey	La, Ta, B	High refraction, low dispersion

glasses have been known for a long time for their beautiful green fluorescence which consists of bands of a position which is only slightly affected by an electrical field. Weyl and Thümen[3] demonstrated that the uranyl group UO_2^{2+}, in which oxygen shields the uranium, is the carrier of this green fluorescence while uranous (U^{4+}) and uranate (UO_4^{2-}) glasses do not fluoresce.

The increase in silica in an alkali silicate glass leads to an increase of the intensity of the UO_2^{2+} fluorescence, according to Rodriguez, Parmelee, and Badger.[6] (Fig. 3.)

The curve can be interpreted as an indication of the formation of a stronger network.

Similarly, the intensity of the UO_2^{2+} fluorescence of glasses containing alkali can be interpreted. (Fig. 4.)

Among the monovalent modifying ions Li^+, Na^+, and K^+, the first has the strongest perturbing action because the charge +1 is concentrated in a small volume. Lithium therefore is capable of the greatest disruption of the $Si\text{-}O_4$ network causing the largest decrease in the intensity of the fluorescence of uranyl.

The fluorescence of some rare earth elements is very intensive in part because the processes of absorption and emission of light are performed by electrons of inner shells. The blue fluorescence of cerium glasses, like Crookes eyeglasses, is best known. A samarium glass lends itself to demonstration because of its attractive orange-red fluorescence.

FIG. 3. Intensity of UO_2^{2+} fluorescence *versus* SiO_2 content in an alkali silicate glass, according to Rodriguez, Parmelee, and Badger.

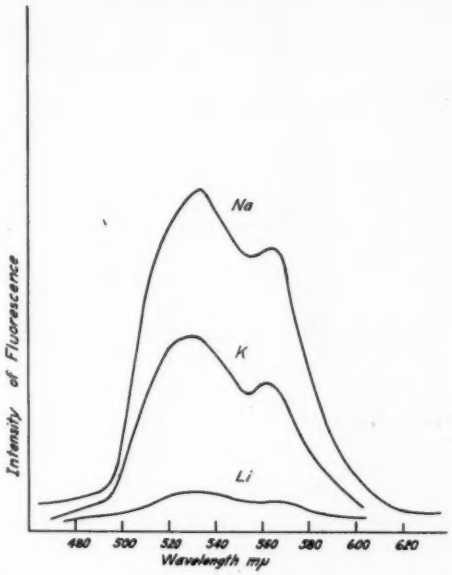


FIG. 4. Comparison of the intensity of the UO_2^{2+} fluorescence in potassium, sodium, and lithium silicate glasses.

Europium ions have been applied to the structural interpretation of the glassy condition or vitreous state. Even though the fluorescence is less sensitive to external influences, europium is useful for the purpose of structural analysis because its bands and lines are well defined.[7] Not all of the electronic transitions are equally suitable to compute the symmetry and intensity of outer electric fields from the shift to higher or lower frequency.

A relatively simple case is represented by the 5790A line which cannot be split into two or more lines by an electrical field.

Tomaschek and Deutschbein[8] found that this wave-length is decreased as the Eu^{3+} ion is

[6] A. R. Rodriguez, C. W. Parmelee, and A. E. Badger, J. Am. Cer. Soc. **26**, 137–150 (1943).

[7] H. Gobrecht, Ann. d. Physik **28**, 673–700 (1937); abstr. C. A. **31**, 947[8].

[8] R. Tomaschek and O. Deutschbein, Glastech. Ber. [5] **16**, 155–161 (1938).

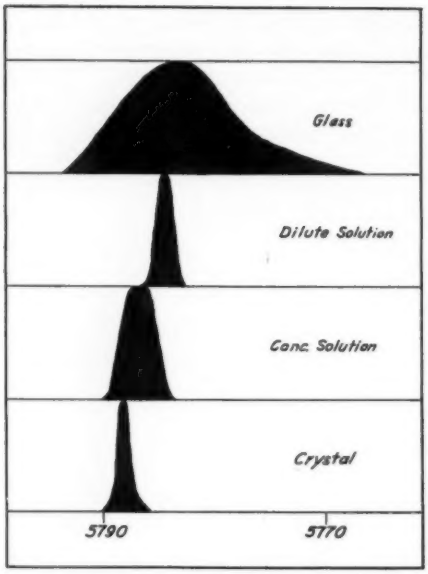


FIG. 5. Position and width of 5790A fluorescence band of europium in various media, according to Tomaschek and Deutschbein.

brought from crystals into concentrated solutions, diluted solutions and, finally, into glass. (Fig. 5.)

This decrease of wave-length was interpreted as a decrease of the surrounding electrical field. In this respect it is significant that the glass does not rank between crystal and solution. The electrical field of the environment of the fluorescence center is the smallest in glass.

In a way this can be considered a consequence of the weaker, or modifying, cations being fixed in rigid positions more remote from their neighbors in contrast to their mobility in a solution. The europium, occasionally taking the position of a cation, signalizes this fact. The wave-length of the band increases further as one proceeds to more concentrated solutions. However, the width of the band of solutions increases in the reverse order because of disturbed symmetry.

5. MANGANESE GLASSES

Divalent manganese, Mn^{2+}, as an activator of crystals or glasses, produces fluorescence colors changing from green to red in various media, these changes being much easier to observe than the fine displacements of lines or their intensity variations. As in crystals, two bands exist in glass, one in the short wave region, the other in the longer wave ranges. It is mainly the latter which is responsible for the green to red colors observed.

Manganese generally occurs in glass as Mn^{3+} and Mn^{2+}. The former, Mn^{3+}, imparts strong purple color and no fluorescence. Mn^{2+} alone fluoresces from red to green depending on the base glass composition. In the small concentrations required for promoting this fluorescence (0.5–2.0 percent) it imparts no noticeable color to the glass. The simultaneous presence of Mn^{3+} covers the fluorescence of Mn^{2+}, but by the use of reducing agents in the melting process the manganese can be easily reduced to Mn^{2+}.

Linwood and Weyl[9] attributed the green fluorescence to Mn^{2+} in a glass providing chiefly the environment of interlinked glass forming units like $(SiO_4)^{-4}$ and the red band to the stronger disturbance caused by a loosened structure. This red fluorescence has been found more affected by the thermal vibrations because of the less rigid bonds in its surrounding.

Some of the color shifts observed by these authors are summarized in Table IV.

In borate and silicate glasses SiO_2 caused a shift of the color toward green in confirmation of the concept that a stronger structure decreases the disturbance of the fluorescence center. The addition of titanium and aluminum oxide in silicate glasses acted in the same direction proving the "network forming" nature of these oxides which are incapable of forming a glass by themselves. Sodium in silicates, borates, and

TABLE IV.

Mn^{++}	Green	Red
Silicates	Si←	→Na
	Ti←	
	Al←	
Borates	Si←	→Na
Phosphates	Zn←	→Na
Coordination		→
Temperature change:	slight	rapid
Concentration change:	rapid	slight

Fluorescence of Mn^{++} glass from data by Linwood and Weyl

[9] S. H. Linwood and W. A. Weyl, J. Opt. Soc. Am. **32**, 443–453 (1942).

phosphates conversely favored the red color indicating its loosening or "modifying" character. Zinc, in the phosphate glasses investigated, revealed itself as a strengthening or glass forming addition by the shift to green imparted to the fluorescence of the glass.

In the course of a study of optical glasses and filter glasses the writer has melted several glasses in which molecular replacements were made on the basis of the systems:

(1) Na_2O, $4SiO_2$,
(2) Na_2O, $2SiO_2$,
(3) Al_2O_3, Na_2O, P_2O_5,
(4) Al_2O_3, BaO, ZnO, P_2O_5.

(A) The Influence of the Network Former

When the ratio SiO_2 to Na_2O is decreased the fluorescence color changes from green to yellow. This shift indicates a loosening of the structure and a stronger disturbance of the fluorescence center.

The color change is represented in Table V, Chart "A." Pigments of spectral green and spectral red were mixed in various ratios and a series of twenty just perceptible steps was arranged. The series of colors is represented by twenty squares. The fluorescent colors of the glasses were compared to these colors and noted by inserting the chemical symbol of the change made in a bar indicating the respective square.

(B) The Series K+—Na+—Li+

The fluorescence color is shifted from blue-green to yellow if the modifiers are exchanged in a series K+−Na+−Li+. The loosening of the glass structure by the Li+ leads easily to devitrification, i.e., crystallization of the glass. Therefore, only one-half of the sodium was replaced by K+ and Li+.

The composition of the glasses of this series thus was:

Type	
K+	$\frac{1}{2}K_2O$, $\frac{1}{2}Na_2O$, $4SiO_2$,
Na+	Na_2O, $4SiO_2$,
Li+	$\frac{1}{2}Na_2O$, $\frac{1}{2}Li_2O$, $4SiO_2$.

The color change is represented in Chart "B" of Table V.

The field strengths of the modifying ions K+, Na+, and Li+ can be expressed as the ratios of their charges over the square of their apparent radii:

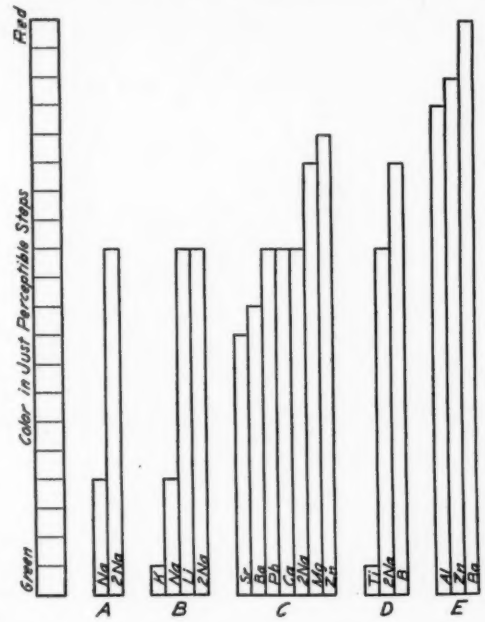
K+	0.26,
Na+	0.38,
Li+	0.46.

The increase in wave-length of the fluorescence of the three glasses in this order may be explained again by the increasing disturbance caused by the structure being loosened by modifiers of increasing field strength.

(C) The Divalent Elements

If in the glass $Na_2O \cdot 2SiO_2$ half of the Na_2O was replaced by RO, where R represents one of the divalent ions Zn^{2+}, Mg^{2+}, Ca^{2+}, Ba^{2+}, Sr^{2+}, and Pb^{2+} the fluorescence remained practically unchanged. It can be learned that in a sodium silicate glass of this composition the loosening effect of the divalent ions was about that of the

TABLE V. Fluorescent color of Mn^{2+} in systematically varied base glasses. The scale on the left represents twenty just perceptible colors between green and red. The length of the bars indicates to which of these colors the fluorescence of the glasses compares. The characteristic substitution is indicated by the insertion of the chemical symbols of the respective chemical elements. Looser structures in general correspond to a color shift towards red.



sodium, with Mg^{2+} and Zn^{2+} being a little more effective, and Sr^{2+} and Ba^{2+} (Chart "C" of Table V), a little less. The calculated field strengths of these ions: Ba^{2+} 0.5, Sr^{2+} 0.56, Ca^{2+} 0.66, and Mg^{2+} 0.9 are higher than those

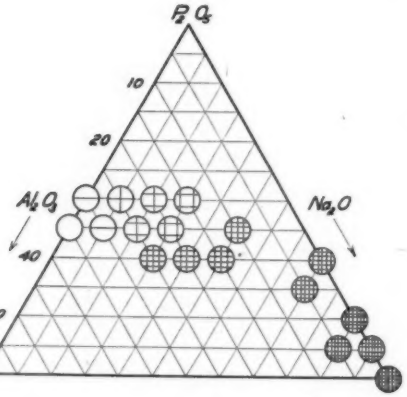


FIG. 6. Fluorescence color of Mn^{2+} in glasses of the ternary system Al_2O_3, P_2O_5, Na_2O. Increasing red hues are symbolized by cross-hatching.

of the monovalent ions. (An estimate of Zn^{2+} or Pb^{2+} seems futile because of the different arrangement of their electrons.)

From the fact that the disturbance is only of the order of that of the monovalent ions, it can be concluded that the divalent ions are capable of a considerable strengthening effect of their own.

It can be seen from the color "signals" that the substitution for Na^+ of Ba^{2+}, Sr^{2+}, Pb^{2+}, and even K^+ acts much more like an addition of SiO_2 to the glass than one would expect on the basis of the distinction between glass formers and modifiers.

In phosphate glasses, where, because of the higher ratio O:P (2.5 in P_2O_5),[10,11] many oxygen ions are not shared by two glass forming cations P^{5+}, a loose structure results, as evidenced by their reddish fluorescence colors. In these glasses the addition of zinc, like Al_2O_3, acts in a strengthening manner on the network and the color shifts from red to yellow (Chart "E," Table V).

In a ternary diagram (Fig. 6) representing the colors of the glasses of the field $Al_2O_3-Na_2O-P_2O_5$ the effect of Al_2O_3 as a strengthener and Na_2O as a loosener of the P_2O_5 network can be demonstrated by the red color changing toward yellow from the Na_2O corner both toward Al_2O_3 and P_2O_5; from P_2O_5 to Al_2O_3 the red content also decreases.

Some other glasses containing Ti and B are listed in Chart "D" of Table V. TiO_2 replacing Na_2O in the Na_2O, $2SiO_2$ series shifts the color toward green; B_2O_3 replacing SiO_2 shifts the color toward red (not represented on chart); B_2O_3 replacing Na_2O did not affect it much. A thorough discussion of these relations would require further experimental work but the character of TiO_2 as a potential glass former and the relative looseness of B_2O_3 are well known to the glass technologist.

Data of this type are used in comparing the structure of unknown optical glasses and in studying the variation of color imparted by the same colorants in filters of new base glass composition.

7. THE IDENTIFICATION OF OPTICAL GLASSES

Many glasses showing weaker or hardly perceptible fluorescence in a mercury vapor lamp of the usual type were found to exhibit a more intense fluorescence in a sterilizing lamp. This type of lamp, emitting shorter wave-lengths of the uv, is now readily available and was applied successfully to the identification of optical glasses.

The identification is based on three incidental variations in the available types of optical glass: (1) The presence or absence of lead. (2) The presence or absence of antimony. (3) The varia-

TABLE VI. Identification of optical glasses.

Example*	Fluorescence
Barium crown	Dark purples; with PbO bluer
Borosilicates	Brilliant purples; with more Sb_2O_3 more purple
Crowns, with Sb	Purplish whites; with more Sb_2O_3 more purple
Crowns, without Sb	White to greenish
Light flints	Brilliant blues
Medium flints	Blue, less purplish, more greenish
Heavy flints	Dull green
Manganese glasses	Greenish yellow

* It should be understood that the fluorescence of these glasses represents examples only. The fluorescence depends less on the optical type, than on incidental or deliberate additions of PbO, As_2O_3, Fe_2O_3, Sb_2O_3. Accordingly a standard scale must be set up in each shop depending on the brand of optical glass used.

[10] N. J. Kreidl and W. A. Weyl, J. Am. Cer. Soc. [11] **24**, 372–78 (1941).

[11] N. J. Kreidl and W. A. Weyl, Glass Ind. **23**, 335–8, 357, 384–6, 402–3, 426–8, 441, 465–7, 480 (1942).

tion of the antimony fluorescence in soda lime, barium, and borosilicate glasses.

(a) Lead

The presence of small amounts of lead is revealed by a blue fluorescence so brilliant that the room may remain illuminated while a crown flint or light flint glass is inspected.

With increasing lead content the color becomes less intense and wanders from purple to blue and green.

Ordinary flints and crown flints can be distinguished easily from each other and from crown glass. Even more than a dozen different flint glasses could be arranged in a series that was found in fair agreement with their lead content (Fig. 7.)

(b) Antimony

Antimony in concentrations of the order of $\frac{1}{2}$ to 1 percent produces purple fluorescence. If the antimony content decreased below $\frac{1}{2}$ percent, the fluorescence shifts through white to green. In this way, the Bausch and Lomb borosilicate crowns, BSC-1 and BSC-2, could be easily identified.

(c) Base Glass

In borosilicate crowns the color is more saturated. In barium crowns the color is less intense and more bluish, particularly in those containing some lead.

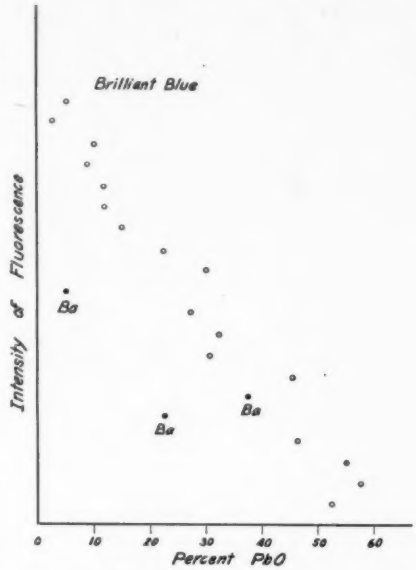


FIG. 7. Intensity of fluorescence of commercial flint glasses.

This influence of the base glass composition permits the distinction of quite a number of commercial crowns, barium crowns, and borosilicate crowns (Table VI). For demonstration purposes an exhibit of optical glasses, a sterilizing lamp, and a suitable Corex type of filter may be assembled into an attractive display.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 4 APRIL, 1945

A Method for Determining the Chromaticity of Fluorescent Material

W. H. Byler and C. C. Carroll
United States Radium Corporation, New York, New York
(Received December 27, 1944)

WITH the increasingly widespread application of fluorescent materials and with the growing recognition of the significance of color limits in special applications, the need for an adequate method of measuring the chromaticity becomes urgent. We are concerned here principally with applications where the exciting source is either ultraviolet light or radioactive substances, although the method can be applied to any type of excitation and all ranges of luminance. In the past, the chromaticities of low luminance materials have been determined by making over-all visual comparisons with incandescent lamp-filter combinations. But, the eye is not capable of detecting differences in dominant wave-length, excitation purity, and luminance independently. There is need for a method based upon determination of spectral energy distribution.

Some of the fluorescent materials of interest have such low luminance that the choice of type of radiant energy detector is limited to the photographic emulsion. Others can be measured with the visual spectrophotometer, but this is too tedious for routine work. Instruments employing the photo-cell which are available at present are not designed for such measurements. In fact, present photo-cells are probably not sensitive enough for this problem. The spectrographic method has the desired sensitivity and avoids the principal difficulties of the visual method. The usual objections to the photographic method as applied to ordinary photometry are not valid here. It will be shown later how the different film variables are kept within acceptable limits.

The core of this method is the procedure for calibrating the Tri-X panchromatic film. This was done by comparing film density curves of two phosphors with their relative energy curves obtained with a spectrophotometer. Having once found the factors required to convert film density to relative energy, spectral energy curves for unknown fluorescent colors can be readily derived from their spectrograms.

Two phosphors were chosen, a green and an orange, whose combined emission covers the range from 420 to 650 millimicrons with sufficient luminance to allow measurement of the energy distribution with a visual spectrophotometer. A thin layer of the phosphor was spread on a glass slide and mounted in front of the comparison prism of a Bausch and Lomb spectrophotometer. A P7 Sylvania Blacklite lamp was placed near the slide and readings taken at intervals of 5 millimicrons. The orange phosphor was prepared to be as light-stable as possible, consistent with a workable level of brightness. In order to minimize possible error owing to discoloration of the sample during the determination, separate runs were made starting from the blue and red ends of the spectrum and at the peak. Before starting determinations, the added precaution of heating the slides above 100°C to drive off moisture was taken. It was necessary to place a neutral filter in the beam of the spectrophotometer lamp for work with the orange phosphor. Three curves were run on each of the two phosphors and a mean curve drawn for each of them. The energy distribution of the spectrophotometer lamps was determined by measuring the color temperature using an optical pyrometer calibrated against a lamp which had been calibrated at the Bureau of Standards. An energy curve for the lamps was then drawn by interpolating between known curves. Combining the spectrophotometer curves for the phosphors with this energy curve, relative energy curves for the phosphors were drawn (Fig. 1).

Spectrograms for the phosphors were obtained with a Cenco grating spectrograph. This instrument has a dispersion of about 1.6 millimicrons per mm. A mercury arc reference spectrum was recorded on each film, and a slit width of 0.75 mm was used for phosphor exposures. A thin layer of phosphor was spread on a glass plate and mounted against the spectrograph. A Corning No. 587 filter, which cuts out below 310 mu, was mounted between the slide and the P7 lamp. This was found

to be necessary because radiation from the bare P7 lamp darkens some phosphors during exposure. The film was developed according to recommendations of the manufacturer. Under these conditions, the green phosphor yielded a peak density above fog of 0.765 for a 45-min. exposure and the orange phosphor 0.665 for a 225-min. exposure.

In determining the film calibration curve, the relative energy and film density curves were drawn so that the factor would be unity at 555 millimicrons for each phosphor (Fig. 1). This is legitimate since we are interested in relative energy curves. The final curve is shown in Fig. 2. Ordinates on this curve represent the ratios of energy to density values at corresponding points on the curves of Fig. 1.

Measures taken to minimize some of the errors in the spectrophotometer work have been mentioned. The difficulty of making matches in the deep blue should be pointed out. This work should be done with a well-rested and dark-adapted eye. However, the spectrophotometer curves can be determined for the chosen phosphors once and for all, and as many curves as needed to achieve the desired accuracy can be combined. In this calibration, only the points from 430 millimicrons down failed to fall very near a smooth curve. Having once determined the relative energy curves for the phosphors, it is simple to calibrate new film. If desired, each film used for a determination can be calibrated, since a number of exposures can be made on a single film.

It should be noted that this method involves the assumption that film density is proportional to energy density. Of course, this is not strictly

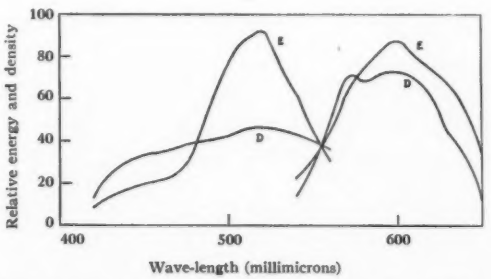


FIG. 1. Spectral distribution and film density curves of two phosphors. *E*: energy; *D*: density.

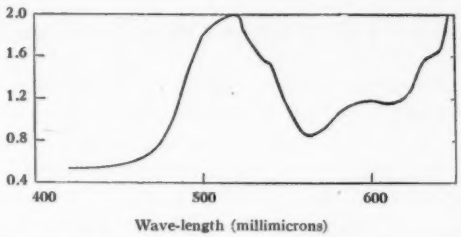


FIG. 2. Tri-X panchromatic film calibration curve. Film density×ordinate equals relative energy.

true since it assumes validity of the reciprocity law, it assumes a straight line relation between density and exposure, and it assumes constant slope of the density-exposure curve at all wave-lengths. But, we have found that failure of the reciprocity law does not introduce a significant error, and steps can be taken to keep errors due to the other assumptions at an acceptable level.

It is true that failure of the reciprocity law is quite significant in ordinary photometry, for in that case a 10 percent failure of the law would result in a 10 percent error in the final result. According to Mees,[1] reciprocity law failure is not a function of wave-length. Therefore, we are not concerned here with the positions of the high and low portions of the curves. We need consider only the low intensity component of reciprocity law failure. The effect of this failure is to make the low density portions of the curves lower than they should be relative to the high density portions. The error introduced by a slight inaccuracy in level of this portion of the curve, which makes only a small contribution anyway to the final result, is insignificant. Justification for neglecting this error can best be shown by an example. In calculating the film calibration factors of Fig. 2, we used overlapping values from the two phosphors at 5-millimicron intervals from 540 to 560 millimicrons (see Fig. 1). Comparing the factors obtained from the two phosphors, we find the greatest difference at 540 millimicrons where it amounted to 9 percent of the mean value. At this wave-length, we had an actual film density ratio of 5.35 for the two phosphors; and since the exposure time ratio was 5.0, the irradiance ratio was of the order of 25. This is a rather wide range for the purpose of this discussion. Taking the

[1] C. E. K. Mees, *The Theory of the Photographic Process* (The Macmillan Company, New York, 1942).

complete density curve for the orange phosphor, which yielded readable values all the way to 420 millimicrons, and applying the calibration factors of Fig. 2, we calculated the chromaticity and found the dominant wave-length to be 585.5 millimicrons and the excitation purity 95 percent. Now, let us assume an extreme 10 percent failure of the reciprocity law for the portion of the curve from 550 down to 420 millimicrons. Adding 10 percent to all points of the curve below 550 millimicrons, we find a dominant wave-length of 585 millimicrons and excitation purity 94.5 percent. We have made direct tests of the effect of reciprocity law failure by varying lamp distance and by using lamps of different ultraviolet light output. These tests show the effect to be unimportant.

As for the assumption of a linear relation between density and exposure, this is nearly enough true for the degree of accuracy required provided net peak density is always kept below 1.0.

Differences of slope of the density-exposure curve at different wave-lengths are small, but appreciable error can arise from this source if peak density is too far from that used in the calibration. From the peak densities given above and the relative density curves of Fig. 1, it can be seen that the densities employed in the calibration were quite uniform across the spectrum except for the extreme blue and red. We find that peak densities in the blue to yellow range can depart widely from these values without much effect on calculated chromaticity, but it is recommended that all peak densities be kept within 20 percent of those employed in the calibration.

Having obtained the calibration curve, then, the procedure for determining the chromaticity of unknown fluorescent colors is to obtain a spectrogram and read the densities above fog at as many points as desired, and multiply these readings by ordinates at corresponding wave-lengths on the calibration curve. From the curve thus obtained, the trichromatic coefficients may be calculated by the selected ordinate method using a table of selected ordinates for evaluation of radiant energy.[2,3] Thirty ordinates are sufficient for most phosphors. The chromaticities of a large number of zinc and zinc-cadmium sulfides have been determined in this laboratory by this method. We have found that we can reproduce results well within the limits of the ability of the eye to distinguish wave-length differences.[2]

[2] Committee on Colorimetry, J. Opt. Soc. Am. **34**, 633 (1944).
[3] A. C. Hardy, *Handbook of Colorimetry* (Massachusetts Institute of Technology Press, Cambridge, 1936).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 4 APRIL, 1945

Dark Adaptation Following Light Adaptation to Red and White Lights* [1]

SELIG HECHT AND YUN HSIA[2]
Laboratory of Biophysics, Columbia University, New York, New York
(Received January 20, 1945)

I. PHOTOPIC AND SCOTOPIC VISION

THE double structure of the vertebrate retina involves separate receptors for night vision and for day vision. In man this is responsible for the classical Purkinje phenomenon (Purkinje, 1825) in terms of which the relative luminosity of the spectrum varies with the brightness (Koenig and Ritter, 1891), and finds its quantitative expression in the scotopic and photopic luminosity curves shown in Fig. 1. The duality of human vision was first formulated by Schultze (1866) and later independently by Parinaud (1881) and by von Kries (1895). Since then it has served as theoretical matrix for a host of visual phenomena (cf. summary by Hecht, 1937).

One of its more practical consequences has been the use of red light for achieving and preserving dark adaptation under conditions permitting the eyes to be occupied with tasks requiring relatively high visual acuity and brightness (Miles, 1943). This has been common practice in fluoroscopy rooms, and has come to have extensive use in the armed forces.

Recently Lowry (1943) has taken exception to such use of red light and has presented data to show that the important factor in determining the time necessary for dark adaptation is not the spectral composition but the brightness of the light to which the eye has been adapted. In order to resolve this contradiction we have re-examined the problem theoretically and experimentally.[3]

II. REASON FOR RED LIGHT

In explaining the use of red light for achieving dark adaptation, it is usually said that the retinal rods which mediate scotopic vision are practically insensitive to red light by comparison with the cones which mediate photopic vision. As evidence, one generally refers to the scotopic and photopic luminosity curves of Fig. 1 from which such a conclusion might seem reasonable.

This, however, is not correct. Each of the luminosity curves is drawn independently, and describes the comparative effectiveness of the spectrum for its own system only. The relative luminosities of the two systems are not given, because the maximum of each curve is arbitrarily put at 100.

Moreover, it is not even true that, compared to the cones, the rods are practically insensitive to red light. Actually in peripheral vision, the rods are slightly more sensitive to red light than the cones. This becomes clear from the comparison of the two luminosity curves on a real energy basis such as is shown in Fig. 2 (Hecht, Shlaer, and Smith, 1935). At any wave-length, as the energy is increased, it is first recognized

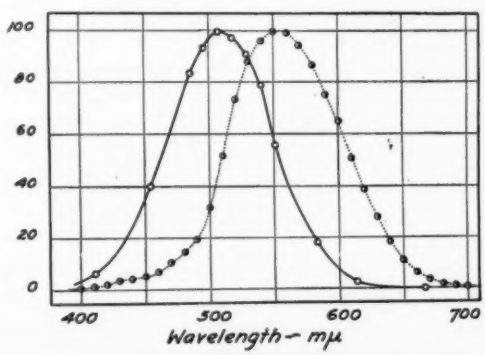


FIG. 1. Luminosity curves for scotopic (rod) vision and for photopic (cone) vision. The scotopic data are the measurements of Hecht and Williams (1922); the photopic data are those of Gibson and Tyndall (1923). Since the maxima are arbitrarily set at 100, these curves give no information about the relative sensitivity of rods and cones for red, or for any other part of the spectrum.

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] A report of these findings was given to the Optical Society of America at its meeting on October 21, 1944 in New York.

[2] Fellow of the China Institute of Physiology and Psychology.

[3] After our work was completed and reported, a paper appeared by Rowland and Sloan (1944) in which they demonstrated experimentally the value of red over white light for preserving dark adaptation. In addition, they made an approximate determination of the numerical advantage of red over white. Our own approach is sufficiently different to constitute an independent confirmation of the theory and practice of red light pre-adaptation.

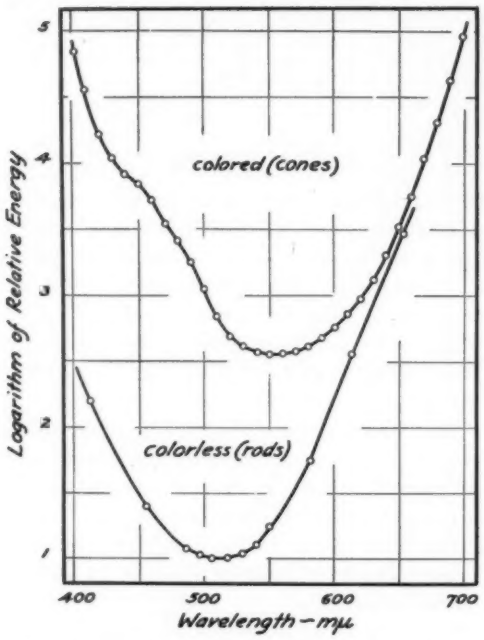


FIG. 2. Spectrum sensibility curves for rod and cone vision on a real energy basis. The data for the separate curves are from the same sources as in Fig. 1. The position of the two curves on the ordinates corresponds to the fact that after complete dark adaptation, any region of the retina outside the fovea sees red light of 650 mμ as colorless at the threshold, and as colored only above the threshold. The precise energy increment above the threshold for the appearance of color (cone function) varies for different parts of the retina; in the parafovea it lies between 0.1 and 1.0 log unit.

by the rods as colorless light, and then at higher energies by the cones as colored light. The energy difference between the two systems at different wave-lengths is the photochromatic interval of Charpentier (1880) and Parinaud (1898) and shows up in many visual functions such as dark adaptation, instantaneous threshold, flicker, intensity discrimination, and visual acuity (Hecht, 1937).

The fact that the rods and cones have practically identical sensibilities for red light is confirmed by the measurements of Mandelbaum (1941) who found that from fovea to extreme periphery, covering all variations of rod and cone population, the sensitivity of the retina to red light remains practically constant. Thus since the rods in scotopic vision are at least as sensitive as the cones and usually more so, this cannot be the reason for using red light or red transmitting glasses for achieving dark adaptation.

The real reason is somewhat more subtle, and depends on the relative positions of the two luminosity curves on the wave-length scale of the spectrum. What is required is the evaluation of the relative effectiveness for the two systems of different lights of specific energy and spectral composition.

For simplicity assume a light which emits equal energy throughout the visible spectrum, and a filter which has a sharp cut-off, so that it transmits light only beyond 620 mμ. The visual effectiveness of the whole white light for the cone system is proportional to the area under the photopic visibility curve in Fig. 1, while the effectiveness of the transmitted red light alone is the area under the curve to the right of a vertical line at 620 mμ. The relative brightness of whole light compared to the red light will be the ratio of these two areas. For the photopic curve this ratio is very nearly 10, which means that if for the cones whole white light is to be made equal in brightness to the red light transmitted from it through the red filter, the whole light has to be reduced to 1/10 its intensity by means of a neutral filter or any other device which decreases it evenly along the spectrum.

Now consider the scotopic curve for the rods. Because it is about 50 mμ to the left of the cone curve the ratio of the total area under it to the area beyond 620 mμ will be much greater than 10. Actually it is greater than 100, which means that, for the rods, if whole light is to be made equal in effectiveness to the red light transmitted by the filter, the whole light has to be reduced to less than 1/100 by a neutral device. It follows that when whole white light and red light have been equalized for cone vision, then for rod vision the white light will still be 10 times brighter than the red light. Thus by comparison to white, red light has a tenfold advantage in the dark adaptation of the rods for an equal effectiveness in cone vision.

The ratios here used are approximate. The precise values will depend not only on the photopic and scotopic luminosity curves in Fig. 1 —call these P and S—but on the energy distribution E in the spectrum of the lamp and on the

transmission T of the filter. At any wave-length λ these values will be P_λ, S_λ, E_λ, and T_λ. Thus in the visible spectrum between 400 and 700 mμ the brightness of whole white light W for photopic vision is

$$W_P = \int_{400}^{700} E_\lambda P_\lambda d\lambda, \quad (1)$$

and for the filtered red light R it is

$$R_P = \int_{400}^{700} E_\lambda P_\lambda T_\lambda d\lambda, \quad (2)$$

while for scotopic vision the comparable value for white is

$$W_S = \int_{400}^{700} E_\lambda S_\lambda d\lambda, \quad (3)$$

and for the transmitted red

$$R_S = \int_{400}^{700} E_\lambda S_\lambda T_\lambda d\lambda. \quad (4)$$

For photopic vision the ratio of white to red is $B_P = W_P/R_P$, while for scotopic vision the ratio is $B_S = W_S/R_S$. The relative advantage in rod vision of red over white when both have been equalized in brightness for the cones is given by the ratio B_S/B_P.

The advantage ratio B_S/B_P will depend not only on the luminosity curves in Fig. 1, but on the energy distribution in the spectrum of the source and on the spectral transmission of the red filter. For daylight and for tungsten lamps and for common red cut-off filters, computations show that red has a dark adaptation advantage over white of 10 to 50 times.

These values are subject to variation because of individual differences in visibility curves. In addition, the visibility curves themselves, such as the accepted I.C.I. values or even the original data of Gibson and Tyndall (1923) are open to revision (see summary by Federov, Federova, Plakhov, and Seletzkaya, 1940). Similarly the scotopic luminosity function (Hecht and Williams, 1922) may be too high in the red. However, for present purposes they are adequate.

In his argument Lowry has ignored the existence of all this knowledge of scotopic and photopic vision and has tacitly assumed that two lights matched in brightness with the photopic cone system will be equally effective on the scotopic rod system.

III. RED AND WHITE LIGHT ADAPTATION

Our experiments were conducted in three stages. Our first desire was to make certain that the common experience with dark adaptation following red and white illumination is correct. For this purpose we measured dark adaptation with a standard adaptometer (Hecht and Shlaer, 1938) following light adaptation to white light and to red light of approximately equal photopic brightness.

The white adaptation field was illuminated by the regular 40-watt lamp of the adaptometer. On 115 volts, its color temperature is 2600°K as measured with the Eastman Color Temperature Meter. With neutral filters the brightness of the adaptation field was adjusted to 26.3 millilamberts.

For red light we used the dark-adaptor plastic material made by the Polaroid Corporation. Its

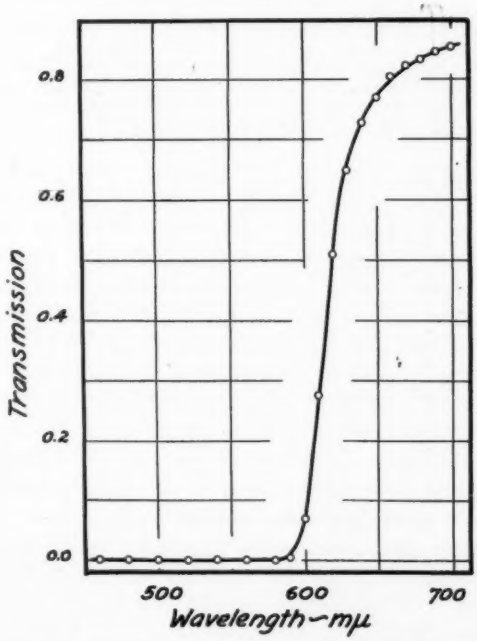


FIG. 3. Transmission of the red plastic material used in these experiments. This is the common filter material supplied by the Polaroid Corporation in the red goggles used for dark adaptation.

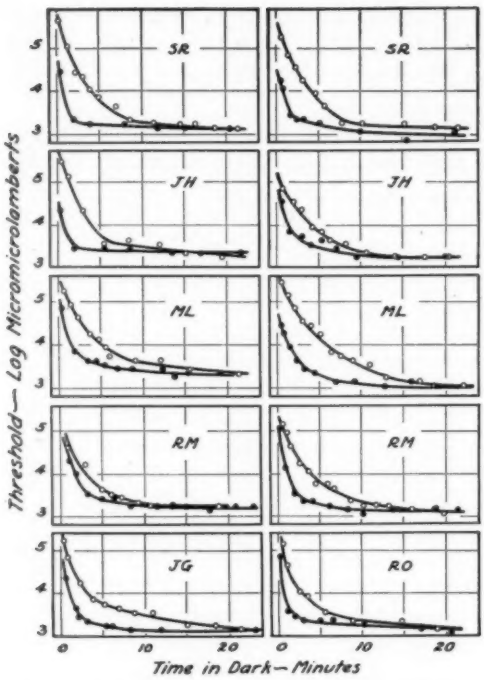


FIG. 4. Individual measurements of dark adaptation following adaptation to white light (clear circles) or to red light (filled circles) of approximately equal photopic brightness. Five subjects were tested twice; two only once. In the left column the red experiment preceded the white; in the right column white preceded red.

spectral transmission is shown in Fig. 3 and was measured with Shlaer's (1938) photoelectric spectrophotometer. From Eqs. (1) and (2) we computed the photopic transmission of the red filter to be 0.129. Visual heterochromic measurement of its transmission made with the Koenig-Martens polarization photometer at high brightness confirmed the computed value. The filter was placed in the path of the adapting light and its photopic brightness adjusted to be 38.9 millilamberts. The brightnesses of the red and white were planned to be approximately equal, but owing to a computational error, discovered after the experiments had been made, they turned out to be different but still of the same order of magnitude. As will be apparent, the error is of no importance.

Each subject was first dark adapted. His right eye was then exposed to the light adapting field for five minutes, after which its dark adaptation was measured using a 3° circular blue field placed 7° off-center exposed for flashes of 0.2 second.

Figure 4 shows the results with seven subjects. Five of them were tested twice, once with red adaptation preceding white and then with white adaptation preceding red. The other two subjects were tested only once.

It is obvious from Fig. 4 as well as from the average data in Fig. 5 that dark adaptation following red light adaptation is much more rapidly attained than following white adaptation. If we choose as a measure of the speed of dark adaptation the time required to reach a level of sensitivity 0.5 log unit above the final threshold, the speed of adaptation subsequent to red is more than three times as great as that subsequent to white. The average time to reach this level after red is 1.9 minutes; after white, it is 6.2 minutes.

IV. LIGHT ADAPTATION AND DARK ADAPTATION

The second series of experiments was designed to clarify the situation further. It is known from previous work (Hecht, Haig, and Chase, 1937) that the speed of dark adaptation varies inversely with the brightness of the preceding light adaptation. The relationship is such that the time in the dark required to reach a given threshold is a sigmoid function of the logarithm of the pre-adaptation brightness. We decided to measure for white light and for red light this relation between pre-adaptation brightness and subsequent dark adaptation.

As before the experiments were made with the Hecht-Shlaer adaptometer with the self-contained adaptation light whose brightness and

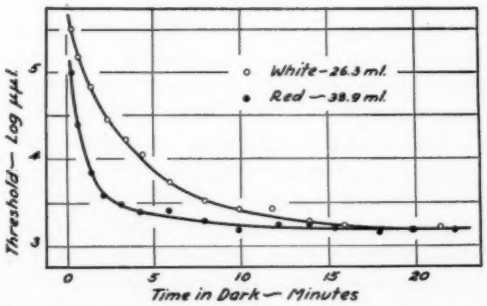


FIG. 5. Average measurements from Fig. 4. It is plain that following white adaptation, dark adaptation (which here is purely rod) is slower than following red adaptation.

spectral composition may be varied by neutral and color filters. The subject was first thoroughly dark adapted. He was then light adapted for 5 minutes to a specified light, after which the course of his dark adaptation was measured with blue light for the same retinal area as before. For light adaptation we used three intensities of white light, and three intensities of red light arranged in approximately equal pairs of white and red. Because of the same error, the red and white lights were not quite the same, but this is of no importance since the two groups of measurements are to be treated independently.

The results obtained with one individual are shown in Fig. 6. At each intensity level for red and for white the measurements were repeated once. They were so unequivocal that we found it unnecessary to use additional subjects. From Fig. 6 it is apparent that the speed of dark adaptation decreases as the intensity of the pre-adaptation increases. Moreover, adaptation following white exposure is always slower than that following red adaptation. The differences in rate of dark adaptation following red and following white pre-adaptation appears to be greater at the higher intensities. However, this may be only because of the scale of plotting the ordinates.

To compare the speed of adaptation one may choose a point on the dark adaptation curve which can be determined with good precision.

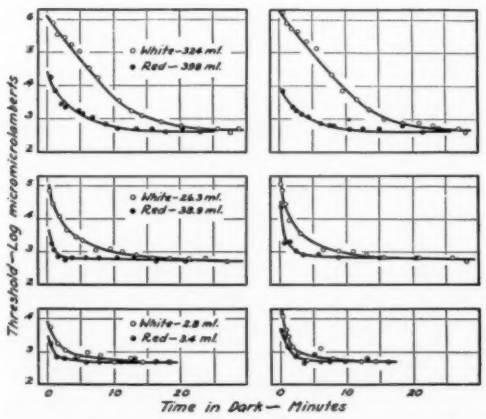


FIG. 6. Dark adaptation following adaptation to white light or red light at three brightness levels. Two runs are shown for each color and brightness for one subject. Regardless of brightness level, dark adaptation is always faster after red light adaptation than after white of comparable brightness.

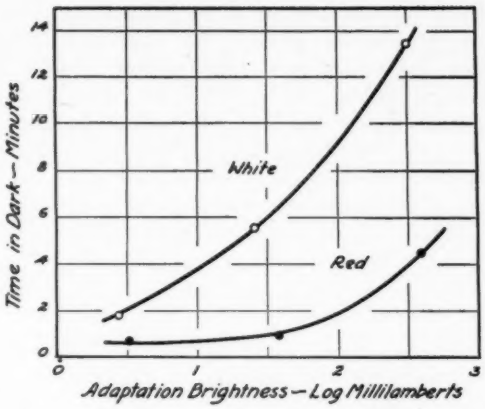


FIG. 7. Relation between the brightness of the preceding light adaptation and the time taken during the subsequent darkness for the eye to reach a threshold of 0.5 log unit above the final threshold. The two curves are roughly parallel, with the red displaced nearly 1.5 log units to the right. This means that to secure the same dark adaptation in the same time, the eye may be exposed to 30 times as much red light as white light.

Such a value is 0.5 log unit above the final threshold because this is approximately where the curves have the maximum inflection.

The average of two such determinations on the data of Fig. 6 are shown in Fig. 7. This shows that for white and for red pre-adaptation the time to reach a given threshold level in the subsequent dark adaptation increases with the intensity of the pre-adaptation. The two curves look like the beginning portions of a sigmoid relationship.

The important point is that the red function is well to the right of the white function. The two curves are approximately parallel and the displacement between them is about 1.5 log unit. This means that to achieve a given degree of dark adaptation one may be adapted to red light which is approximately 30 times brighter than white light. It will be recalled that this is about the expected ratio computed in an early section of this paper.[4]

V. LOWRY'S EXPERIMENT REPEATED

In an effort to reconcile our findings with those described by Lowry we repeated his experiment. We arranged a large sheet of opal glass illu-

[4] Rowland and Sloan found in two experiments that to match the effect of 58 millilamberts of red light they needed between 2.0 and 1.1 millilamberts of white light. This is almost the same ratio as ours.

minated from behind by a lamp whose color temperature was 2800°K. The brightness of the field was adjusted to 4.3 millilamberts with neutral filters. Similarly with the dark adaptor plastic filter and neutral filters we adjusted the field to yield red light of 4.4 millilamberts. These two photopically equal brightnesses were used for light adaptation, and were as near as we could conveniently come with our filters to the 3.15 millilambert field used by Lowry.

Before an experiment the subject was thoroughly dark adapted. He was then light adapted for 3 minutes to either the red or white field. When the light was turned off the speed of his dark adaptation was measured by a method similar to Lowry's, in which we determined the time required to see binocularly a white test field 3° in diameter, 7° below center, set at 0.035 microlambert, and flashed for 1/5 second every second. The brightness of the test field is 4.54 on the log micromicrolambert scale in Figs. 3–5. Lowry's test field was a 12° disk fixated in the center. However, it has been well established (Hecht, Haig, and Wald, 1935) that the threshold of such a large field corresponds to the threshold of its outer edge only. This is because the threshold of the retina for white light decreases rapidly from center to periphery. Thus a small field placed 6° or 7° from the fovea gives about the same threshold as a 12° field fixated centrally.

We measured three subjects three times to each light adaptation. Their results are given in Table I, where each datum is a single determination. Following white light adaptation the average time for all subjects to reach the pre-arranged threshold of 0.035 microlambert is 42 seconds. This is the same order of magnitude as Lowry's results and shows that we were operating under nearly the same conditions. Following red adaptation, however, the measurements show that distinctly less time is required to reach the pre-arranged threshold than before. For all subjects the average is 30 seconds. White light thus causes a 40 percent increase in recovery time over red.

The values of 42 and 30 seconds for adaptation time subsequent to white and red light adaptations are in keeping with the information gained in the preceding section and are indeed to be expected from the curves in Fig. 6. Note that these curves are shallow at low values of pre-adaptation brightness, and steep for high values. For a pre-adaptation of 4 millilamberts red light (log value of 0.6) the recovery time in dark is of the order of a minute for red light. A lengthening of this time by 40 percent calls for a relatively large increase in brightness, approximately a log unit or more.

TABLE I. Recovery after light adaptation. Time in dark to reach fixed threshold of 0.035 microlambert after adaptation to white or red lights of equal photopic brightness, 4.4 millilamberts.

Subject	Y. H.		M. D.		R. W.	
Color	white	red	white	red	white	red
Time in seconds	40	38	29	25	42	30
	41	35	38	25	51	29
	60	41	47	24	42	28
	48	34	—	—	30	27
Average	47	37	38	25	41	29

VI. POSSIBLE EXPLANATION

The mystery of Lowry's results may lie in the second paragraph on page 620 of his paper in which he states that the two adapting fields were adjusted to equality by means of a flicker photometer. The brightness and size of the photometric field are not given. But if the photometric field were larger than 2° or 3°—which is likely—and it were dimmer than the adapting field being measured—which is also likely under the conditions—then Lowry's results are understandable. The photometric measurements will then have been made at low brightness levels at which the Purkinje phenomenon is effective. At such low levels the rods are maximally functional while the cones are only partly functional. And since the rods and cones summate brightness (Koenig and Ritter, 1891) the photometry itself will be mainly scotopic rather than photopic.

If this is the correct diagnosis, then the reason that Lowry got nearly the same results following red and white adaptation is that his red and white lights were of nearly equal stimulating power to the rods and not to the cones. It is also to be noted that there is a small difference even in Lowry's data; the averages differ by 1.1 seconds in the right direction, which for this shallow portion of the curve in Fig. 6 may correspond to a fair photopic brightness difference

between red and white. Lowry's experiments are certainly correct as performed, but they are irrelevant to the phenomenon they were designed to clarify.

SUMMARY

1. Red and white lights, when equated for photopic vision and used for light adaptation, influence subsequent dark adaptation unequally. Following red, dark adaptation is much faster than following white.

2. From the experimental relation between brightness of light adaptation and speed of subsequent dark adaptation, it is found that to produce the same dark adaptation in the same time, red light adaptation may be held 30 times as bright as white.

3. These results conform to theoretical expectation from knowledge of night and day vision, and confirm extensive experience. Experiments which apparently contradict such theory and practice are shown to be in error.

BIBLIOGRAPHY

A. Charpentier, "Le sens de la lumière et le sens des couleurs," *Arch. Ophthal.* Paris, **1**, 48 (1880).

N. T. Fedorov, V. I. Fedorova, A. G. Plakhov, and L. O. Seletzkaya, "A new determination of the relative luminosity curve," J. Phys. U.S.S.R. **3**, 1 (1940).

K. S. Gibson and E. P. T. Tyndall, "The visibility of radiant energy," Sci. Pap. Bur. Stand. **19**, 131 (1923).

S. Hecht, "Rods, cones, and the chemical basis of vision," Physiol. Rev. **17**, 2 (1937).

S. Hecht, C. Haig, and A. M. Chase, "The influence of light adaptation on subsequent dark adaptation of the eye," J. Gen. Physiol. **20**, 831 (1937).

S. Hecht, C. Haig and G. Wald, "The dark adaptation of retinal fields of different size and location," J. Gen. Physiol. **19**, 321 (1935).

S. Hecht and S. Shlaer, "An adaptometer for measuring human dark adaptation," J. Opt. Soc. Am. **28**, 7 (1938).

S. Hecht, S. Shlaer, and E. L. Smith, "Intermittent light stimulation and the duplicity theory of vision," Cold Spring Harbor Symposia on Quantitative Biology **3**, 237 (1935).

S. Hecht and R. E. Williams, "The visibility of monochromatic radiation and the absorption spectrum of visual purple," J. Gen. Physiol. **5**, 1 (1922).

A. Koenig and R. Ritter, "Über den Helligkeitswerthe der Spektralfarben bei verschiedener absoluter Intensität," Zeits. f. Psychol. u. Physiol. Sinnesorgane **7**, 309 (1891).

J. von Kries, "Über die Funktion der Netzhautstäbschen," Zeits. f. Psychol. u. Physiol. Sinnesorgane **9**, 81 (1895).

E. M. Lowry, "Effect of hue on dark adaptation," J. Opt. Soc. Am. **33**, 11 (1943).

J. Mandelbaum, "Dark adaptation," Arch. Ophthal. **26**, 203 (1941).

W. R. Miles, "Red goggles for producing dark adaptation," Federation Proc. **2**, 109 (1943).

H. Parinaud, "L'héméralopie et les fonctions du pourpre visuel," Comptes rendus **93**, 286 (1881).

H. Parinaud, *La Vision* (Octave Doin, Paris, 1898), pp. viii+218.

J. E. Purkinje, Zur Physiol. der Sinne **2**, 109 (1825).

W. M. Rowland and L. L. Sloan, "Relative merits of red and white light of low intensity for adapting the eyes to darkness," J. Opt. Soc. Am. **34**, 10 (1944).

M. Schultze, "Zur Anatomie und Physiologie der Retina," Arch. mikr. Anat. **2**, 175 (1866).

S. Shlaer, "A photoelectric transmission spectrophotometer for the measurement of photosensitive solutions," J. Opt. Soc. Am. **28**, 1 (1938).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 4 APRIL, 1945

Tests for the Detection and Analysis of Color-Blindness

I. The Ishihara Test: An Evaluation

LeGrand H. Hardy, Gertrude Rand, and M. Catherine Rittler
Knapp Memorial Laboratory, Institute of Ophthalmology, New York, New York

(Received February 5, 1945)

The Ishihara test for color-blindness, taken as a representative type of polychromatic tests, fails to give more than a superficial evaluation of the state of an observer's color vision. It can be utilized as a good rough screening test for red-green color blindness, but in so doing most of the plates could be discarded. The results obtained are dependent to a large extent upon the illuminant used, and this fact militates against the deuteranomalous observer. The test gives neither an adequate qualitative (type of defect) nor quantitative (extent of defect) diagnosis, and may give a wrong diagnosis. Results obtained by this test should be discounted unless verified by other modern types of tests.

BECAUSE of their great convenience and apparent simplicity of administration, polychromatic charts for the detection of defective color vision have been widely used in military and industrial fields for many years. Much dissatisfaction has been reported and both the statistics of the Civil Aeronautics Administration and the findings of the Color-Blindness Committee (Army Air Force, U. S. Navy, Civil Aeronautics Administration, and Inter-Society Color Council) indicate that, as they are used at present, polychromatic test plates are only about 50 percent effective in screening out color defectives by the criterion of the test itself.

In an endeavor to discover the reason for this failure, extensive investigations have been undertaken at the Knapp Memorial Laboratories utilizing a large battery of tests on all types of individuals ranging from normals to extreme defectives. Our results indicate that while defects in the tests themselves account for part of the discrepancies, other important sources of error have arisen from: 1. Disregard of the critical importance of illumination. 2. Careless and inefficient administration. 3. Incompetent and invalid interpretation of results. 4. Misapplication of the test to fields for which they were not intended, or the expectation of more from the tests than they were capable of, or designed for, yielding.

The senior author has for years insisted on the critical importance of the illumination used for viewing materials seen by reflected light, claiming that, as they are at present administered, polychromatic tests have become as much tests of the lighting system used as they are tests of the subject's color discriminatory ability.

Likewise he has asserted that results obtained from polychromatic test plates with their low values and chromas and large patterns yielding dim and extensive retinal images, must with great caution be applied to situations requiring the recognition of luminous objects which yield small but bright retinal images. In the absence of time and facilities for carrying out complete colorimetric investigations, at least two types of test should be given an individual before a complete diagnosis is made: one type using reflected light (polychromatic plates, matching chips, 100-hue test or dichotomy test, etc.), the other using luminous signals (anomaloscope, spectroscope, lantern, or threshold test).

This report summarizes a critical evaluation of the Ishihara test plates taken as a representative type of polychromatic plates used for the detection and classification of color defective observers. (The effect of quality of illumination on the results will be reported separately.) With minor changes it can be applied to other plates of the same type (Stilling's, Rabkin's, and the recent American Optical Company's issue of selected plates for the use of the armed services). All have particular merits and defects, and even between different editions of the same test there are to be noted marked, and sometimes critical, variations caused by changes in printing inks, printing papers, printing processes, and age. The use of linseed oil bases for inks is hardly conducive to permanence of printed colors.

Pseudo-iso-chromatic means apparently or falsely equal colors; that is, the colors are trickily chosen to fall within the color diagram zones where the defective observers are most likely to be trapped. Since the apparent color of any object can be changed by modifying the light used to view it, all these plates must be viewed under special external conditions. The Ishihara plates are not particularly critical, and this is one of their defects. Such plates can be made very critical, but in so doing the effect of the illuminant is enhanced.

The Ishihara test for color blindness was devised by Dr. Shinobu Ishihara, professor of ophthalmology, Medical Department, Imperial University of Tokyo and first described by him in 1917 when he was instructor at the Military Medical Academy in Tokyo.[1] Since that time the test has run through eight or more Japanese editions in the course of which some changes have been made in the colors used and in the number of plates in the series, although the basic principles have remained unaltered. The first edition (1917) we have been able to inspect through the courtesy of Dr. Arnold Knapp, professor emeritus of ophthalmology, College of Physicians and Surgeons, New York, who had in his library a complimentary copy of this little-known edition. In this edition the Introduction is less complete than in the later editions, the red-purple pigment is bluer than was used later, and Plates 8 and 9 consist of an entirely different color combination. These plates as first published showed the digits composed of small blue-green disks on a background of small yellow and violet disks. In all later editions the digits are composed of green disks and the background of yellow and red-purple disks. The second edition was copyrighted in 1920* and later editions through the sixth are close copies of this. A third edition was published in 1922 and a fourth in 1924. These were distributed in this country by E. B. Meyrowitz of New York. The first publication on the use of the Ishihara Test in this country was in 1924 by Clark[2] to whom we are indebted for the information that the original plates were destroyed in the Japanese earthquake of 1923. A fifth edition appeared in 1925, at which time C. H. Stoelting Company of Chicago was made the sole distributing agent in the United States, and a sixth edition in 1932. These two editions became widely known and used in this country. Like the earlier editions, they consist of 16 plates, the first plate for demonstration and the detection of color malingering, the 12 succeeding plates for the detection of congenital color-blindness of the red-green type and of monochromasy in literates, and the last three for this detection in illiterates. A seventh edition appeared in 1936, an eighth in 1939, and a ninth in 1940.* These three editions contain 32 plates, the original 16 and 16 new plates of the same type but carrying different patterns. There is also some change in the pigments and a more liberal use of the two-digit plate. The ninth edition was reprinted in London in 1943.

The fifth, seventh, and ninth (British reprint) were the only editions available to us for experimental work and are, therefore, the ones used in the present study and in a study to be reported separately, for a comparison of their value in detecting and analyzing color deficiency.

In the words of the author of the test series, "this series of plates is intended to discover quickly and accurately congenital colour-blindness, the most common form of colour-blindness." "Congenital colour-blindness is of two kinds: Total colour-blindness and red-green blindness, and these have each two forms, complete and incomplete." "Incomplete total colour-blindness is the result of deficiency in the red-green sense and an accompanying weakness of the blue-yellow sense." Ishihara further classifies red-green blindness into red-blindness (the type called by von Kries protanopia) and green-blindness (called by von Kries deuteranopia).[3]

[1] Ishihara, S., *Tests for Colour-Blindness* (Handaya, Tokyo, Hongo Harukicho, 1917).

* For access to this edition we are indebted to Dr. Janet Howell Clark, Professor of Biophysics and Dean of Women of the University of Rochester, and Dr. Anna M. Baetjer, Associate in Physiological Hygiene, School of Hygiene and Public Health, Johns Hopkins University. The book is a complimentary copy belonging to Dr. W. H. Howell, former Professor of Physiology and Director of the School of Hygiene and Public Health, Johns Hopkins University.

[2] Clark, J. H., "The Ishihara test for color blindness," Am. J. Physiol. Opt. **5**, 269–276 (July, 1924).

* For access to the sixth, eighth, and ninth (Japanese) editions we are indebted to Mr. R. Koller of C. H. Stoelting Company.

[3] von Kries, J. "Ueber Farbensysteme," Zeits. f. Psychol. u. Physiol. d. Sinn. **13**, 241–324 (1897).

Ishihara designates red-blindness and green-blindness also as complete and incomplete, but does not describe what he means by this designation. The author makes no claim as to detecting the degree or extent of the color defect, i.e., a qualitative but not a quantitative evaluation is indicated. However, no critical score is given to indicate the dividing line between poor, but normal, color vision and defective color vision.

In constructing his test plates, Ishihara was obviously influenced by Stilling's *Pseudo-Isochromatic Plates for Testing the Color Sense*, but no plates are a direct copy and certain innovations are made.

The plates intended for testing color-blind literates consist of six series, two plates of each series in the fifth edition and four of each in the seventh and ninth editions.

In Series 1, which might be called a "transformation pattern" series, a digit composed in general of blue-red and yellow-red disks varying from 1 to 5 mm in diameter, is seen by the color normal on a background composed of yellow-green disks of these sizes, the disks of both digit and background varying as to chroma and value. To the red-green defective subject, to whom the yellow-green background and the yellow-red disks of the digit are indistinguishable, only that part of the pattern which is formed of the blue-red disks stands out, appearing to him as a blue digit on a yellow background. The pattern is thus transformed for the color defective into a different digit from that seen by the color normal e.g., a three for an eight. This series, as presented in the fifth edition, traps deuteranopes only. (Table I.)

Series 2 also is a transformation series, but with background and pattern color reversed; that is, blue-green and yellow-green disks form the digit on a background of yellow-red. Here the yellow-green section of the pattern is indistinguishable from the background for the color defective and again a blue digit is seen by him on a yellow background. All color defectives fail this series, when tested with the fifth and seventh editions.

Series 3 and 4 are of the "vanishing digit" type used by Stilling. In *Series 3* the digit is formed of yellow-red disks on a background of yellow-green and blue; and in *Series 4* of blue-green disks on a background of blue-red and yellow. These digits are plain to those with normal color vision, but are rarely seen by color defectives unless the defect is slight. This series fails to trap many deuteranomalous and low discrimination observers.

Series 5 contains a "hidden digit" which supposedly is seen only by color defectives, but is too often quite visible to color normals as well. They have little diagnostic value.

Series 6 consists of two digits (e.g. 26), one in red and one in red-purple disks on a background of gray disks of different brightness. Protanopes whose neutral region is in the red, are supposed to see only the red-purple digit (the 6); and deuteranopes, whose neutral region is in the red-purple, only the red digit (the 2). This series helps differentiate dichromasy from trichromasy, and is adequate to classify as to type of defect only deuteranopes.

The remaining plates are for testing illiterates who are instructed to follow "with their eye" a winding line between two terminal *x*'s. In some of the plates this is an easy task for color normals but difficult for the color defectives; in others the path is readily visible to color defectives but not plain to color normals. These plates are difficult to administer and score and, since only literates were included among the subjects used in the present study, they will not be considered further.

EVALUATION OF THE 5TH EDITION OF THE ISHIHARA TEST AS A MEANS OF DETECTING AND CLASSIFYING DEFECTIVE COLOR VISION

The Ishihara test (fifth edition) was included as one test in a battery used in the Knapp Memorial Laboratories to study the cases of defective color vision referred to the laboratories. These cases were not selected on the basis of any statistical survey. They included patients referred to us by ophthalmologists, men who had been rejected by the Armed Forces on the basis of color defect, cases obtained from some of the New York City high schools, from Vanderbilt Clinic, etc. The battery of tests included the following which have been described in the literature:

Ishihara Tests for Colour-Blindness (5th edition)[4]
Rabkin Polychromatic Plates for Testing Colour Vision[5]
Pseudo-Isochromatic Plates for Testing Color Perception, engraved and printed by Beck Engraving Company, Inc., and published and distributed by the American Optical Company[6]
The Farnsworth-Munsell 100-Hue Test[7]
Farnsworth Dichotomy Test[7]
I.S.C.C. Single Judgment Test for Red-Green Discrimination[8]
Nagel Anomaloscope[9]

Several additional tests were used which have not as yet been described in print, notably a series of polychromatic plates and a new form of anomaloscope devised in this laboratory. The results of the entire battery of tests aided in the classification of the type and degree of color defect. Our data thus permit us to evaluate each of the above tests both as a diagnostic or screening medium, and as a medium for the differential classification of type and degree of defect.* The present paper, however, will be confined to the work with the Ishihara test, mention being made of the battery of tests because of its use in the classification of the subjects as to type and extent of defect. Because this battery of tests was so comprehensive, it is understandable why we chose for inclusion the fifth edition of the Ishihara test consisting of 16 plates rather than one of the longer editions. The comparison of results given by the various editions at our disposal was made on a smaller group of subjects, including some of each type of color defective. This work is reported separately.

The entire battery of tests was given to 106 subjects: 23 deuteranomalous and 12 protanomalous cases (anomalous trichromats), 15 deuteranopes and 17 protanopes (dichromats), 7 cases showing low color discrimination in general sufficient to amount to a color defect, and 32 cases who showed low color discrimination to a smaller degree which would not be sufficient to amount, in all probability, to a real color defect. This last group we have called "low normals." This classification of subjects follows the scheme outlined by two of us in "Recent developments in color vision testing."[10] For classifying anomalous trichromasy particular reliance had to be placed on the results of the Anomaloscope, our special Polychromatic Plates and the Farnsworth Dichotomy test; for the diagnosis of low color discrimination and low normal, on these tests and in addition the Farnsworth-Munsell 100-hue test.*

[4] Ishihara, S., *Tests for Colour-Blindness* (Tokyo, Kanehara, 1925), 5th Ed.

[5] Rabkin, E., *Polychromatic Plates for Testing Colour Vision* (Kharkov, State Medical Publishing House of the USSR, 1939), 2nd Ed.

[6] American Optical Company, *Pseudo-Isochromatic Plates for Testing Color Perception* (Beck Engraving Company, Philadelphia and New York, 1940).

[7] Farnsworth, D., "Farnsworth-Munsell 100-hue and dichotomous tests," J. Opt. Soc. Am. **33**, 568–578 (1943).

[8] Hardy, L. H., "Single judgment test for red-green discrimination," J. Opt. Soc. Am. **33**, 512–514 (1943).

[9] Nagel, W. A., "Zwei Apparate für die augenärztliche Funktionsprüfung. Adaptometer und kleines Spektralphotometer (Anomaloskop), Zeits. f. Augenheilk. **17**, 201–222 (1907).

* A test for color-blindness may be a simple diagnostic or screening medium, in which case the purpose is merely to screen out the color defectives from the color normals; or it may be a differential classifying medium, in which case it is designed to determine both the type and extent of the defect. It is our belief that the ideal color-blindness test should serve both functions. (1) For some purposes the simple screening test is adequate and its simplicity of operation and interpretation is an advantage. (2) For such purposes as industrial employment, vocational guidance, and otherwise adapting the environment of the color defective to best serve his practical and esthetic needs, it is important to know how much defect is present and how color stimuli appear to him. There are many vocations, such as automotive wiring, which are open to anomalous trichromats but which cannot be handled by dichromats. Even in this field distinctions must be made. For example, the anomalous trichromat who can distinguish the difference in color between new coded wires viewed under good illumination, is able to work on new motors; but he can not distinguish this difference under flashlight illumination and when the wires have become soiled by usage, and is, therefore, unable to do repair work on the motors. It is often necessary also to know the type of defect and whether or not there is distortion of the brightness value of color stimuli. It would seem that a deuteranope, who has little if any distortion of the brightness aspect of the colors, would be useful in certain types of heterochromatic photometry; while the protanope, who sees the long-wave end of the spectrum as darker and the short-wave end as lighter than normal, could not do this work. Again, a deuteranope is less apt to distinguish between the customary red, green, and amber traffic signals than the protanope, who sees the green signal as bright and the red one as dim. On the other hand, the protanope may find it very difficult to see a red traffic signal during night driving.

[10] Hardy, L. H. and Rand, G., "Recent developments in color vision testing," Am. Acad. Ophth., Outlines of Instructional Courses, 1944, Graduate Lecture, Continuous course No. 6 and 7.

* To those not familiar with the classical terminology of color-blindness and to whom the above reference is not available, the following simple explanation of terms may be of value.

On the Young-Helmholtz theory there are *3 factors in the color vision process:* 1. The red primary, *proton*, or first process; 2. The green primary, *deuteron*, or second process; and 3, the blue primary, *triton*, or third process. Hence we should expect to find *3 types of color vision*. Trichromatic, in which all 3 processes function; dichromatic, in which only two processes function; and monochromatic, in which there is only one differentiated physiological

TABLE I. Analysis of individual plates. Percentage of each type of color defective subject and of the low normals passing each plate of the Ishihara test, fifth edition.

Type of color defective subject	Number of cases	Percentage of each type passing individual plates											
		Series no. 1		2		3		4		5		6	
		Plate no. 2	3	4	5	6	7	8	9	10	11	12	13
Anomalous trichromat		%	%	%	%	%	%	%	%	%	%	%	%
deuteranomalous	23	43	43	0	0	13	22	17	39	13	0	17	30
protanomalous	12	67	67	0	0	0	0	8	8	0	0	8	17
Dichromat													
deuteranope	15	0	7	0	0	0	0	0	0	0	0	0	0
protanope	17	53	59	0	0	0	0	0	0	24	18	0	0
Low discrimination	7	57	71	0	0	14	57	14	43	0	0	43	86
All color defectives	74	42	46	0	0	5	12	8	18	9	4	11	20
Low normals	32	100	100	100	78	87	100	100	94	87	84	97	97

For all work with pigment test material a close approximation to I.C.I. Illuminant *C* was used. The illumination on the test material was about 25 ft.-c. Testing distance in case of polychromatic test plates was about 30 inches. Responses should be immediate (within 2 seconds). Hesitant or studied responses are viewed with suspicion.

The results for the Ishihara test, fifth edition, are summarized in Tables I–IV. They will be discussed under three headings: 1. Analysis of the individual plates; 2. Significance of performance score on the entire test; 3. The differential classification of deuteranopia and protanopia.

process. Trichromatic color vision means all 3 processes are functioning, but since one or more of the processes may function aberrantly there will be as a result *3 types of trichromatic color vision;* 1. Normal; 2. Anomalous, in which one or more of the processes is weak; and 3. Low discrimination in which apparently all processes are weak. The Anomalous Trichromats are designated according to which process is weak as: 1. Protanomalous, or predominantly red weak; 2. Deuteranomalous or predominantly green weak; and 3. Tritanomalous or predominantly blue-yellow weak. Dichromatic color vision means that only two of the processes are functioning and hence is similarly divided into *3 types:* 1. Protanopia, formerly called red blind, 2. Deuteranopia, formerly called green blind, and 3. Tritanopia, formerly called blue-yellow blind. As a mnemonic it may be remembered that there are *3 kinds of abnormal color vision:* 1. Anomalous Trichromasy (*3 types:* protanomalous, deuteranomalous, tritanomalous), 2. Dichromasy (*also 3 types:* protanopia, deuteranopia, tritanopia), and 3. Monochromasy. Protanopia exhibits a shortened red end of the spectrum, a neutral area around 493 mmu (in the blue-green), a second neutral band in the red (at 493 c), and a brightness peak which is shifted toward the violet. Deuteranopia shows a neutral region in the green (about 497 mmu), a second neutral band in the red-purple (at 497 c), and no shift in the luminosity peak. Tritanopia shows a shortened blue end of the spectrum and a neutral band in the yellow-green. The luminosity peak is not shifted. These types, as well as corresponding anomalous trichromatic types, are best diagnosed and classified by anomaloscope tests.

1. Analysis of the Individual Plates

In Table I is shown for each type of color defective subject and for the low normals the percentage of cases who pass the individual Ishihara plates, designated by both series number and plate number. The first five horizontal rows of figures show these percentages for each type of color defective and demonstrate the value of the individual plates for trapping each type. The bottom two rows present for comparison the percentages for the color defectives as a group and the low normals as a group. This comparison shows the value of each plate as a diagnostic or screening medium.

In this table it is seen that the plates of *Series 2* are the only ones that are failed by all the color defectives tested. Plate 4 of this series is passed by all the color normals; Plate 5 by only 78 percent of the color normals. Plate 4 is therefore unique in that all the color defectives fail it and all the color normals pass it. *Judged by our observers, then, this plate when properly administered furnishes an ideal screening or diagnostic test;* used with Mazda light, as will be shown separately, its efficiency for deuteranomalous cases drops from 100 percent to 43 percent. The remaining plates are much less successful. For example, the plates of *Series 3* serve to trap deuteranopes, protanopes and protanomalous cases, but not all deuteranomalous and low discrimination subjects; the plates of *Series 4 and 6* trap deuteranopes and protanopes but not all deuteranomalous, protanomalous and low discrimination subjects; the plates of *Series 5* trap deuteranopes and protanomalous and low discrimination cases, but not all the deutera-

nomalous or the protanopes. The plates of *Series 1* are the least effective, except for deuteranopes. They are particularly ineffective for the protanopic type of subject, who tends to see the digit (which is composed of dominantly reddish disks) as a dark figure and the background (which is composed of greenish disks) as light. For this reason more than 50 percent of protanopes and 67 percent of protanomalous subjects pass these plates. Protanopes pass the plates of *Series 5*, too, in a higher percentage than would be expected; that is, they fail to see the "hidden digit." It is because of their success with plates of *Series 1 and 5* that protanopes achieve, on the average, a higher performance score on the entire test than deuteranopes and nearly the same score as protanomalous subjects (Table II).

2. Significance of Performance Score on Entire Test

Table II shows the performance scores obtained on the entire Ishihara test. In this table are given for each type of color defective and for the low normals (a) the average score achieved by each group; (b) the median or middle score for each group; and (c) the range of scores from lowest to highest achieved by the individuals within each group. A score of 100 indicates that correct responses were given to all 12 plates; a score of 0 that no correct responses were given; and scores between 0 and 100 indicate the percentage of the 12 plates to which correct responses were given. As in Table I, the first five horizontal rows of figures give the data for each type of color defective, and the bottom two rows for the color defectives as a group and the low normals as a group.

In Table II it is seen that (a) the color defectives as a group have an average performance score of 15 and the low normals of 94; (b) deuteranomalous subjects have an average score of 20 and deuteranopes of 1; and (c) protanomalous subjects have an average score of 17 and protanopes of 15. On the average, then, the Ishihara test would seem on the basis of performance score to separate color defectives from color normals and deuteranomalous subjects from deuteranopes. However, for the purpose of screening or classifying it is not the average score achieved by a group that is important but the score achieved by the individual. This is shown in the last column of Table II where for each type of color defective is given the range of scores from lowest to highest achieved by the individuals having that type of defect. This column shows clearly the overlapping of scores between the various types of color defectives and thus demonstrates the *impossibility of differentiating type of color defect on the basis of score alone.* There is, however, no overlapping of scores between the group of color defectives taken as a whole and the group of low normals, the lowest score of the latter group being 75 and the highest score of the former group being 58. In round numbers, then, it would seem that a *critical score of 60 might be accepted as the dividing line between color defective performance and normal performance on the Ishihara test.*

TABLE II. Significance of performance score. Showing for each type of color defective subject and for the low normals (a) the average performance score achieved by each group; (b) the median or middle score for each group; and (c) the range of scores from lowest to highest achieved by the individuals within each group. *Performance score is percentage of the 12 plates to which correct responses were given.*

Type of color defective subject	Number of cases	Average score	Median score	Range of scores
Anomalous trichromat				
deuteranomalous	23	20	17	0– 58
protanomalous	12	15	17	0– 33
Dichromat				
deuteranope	15	1	0	0– 8
protanope	17	13	17	0– 33
Low discrimination	7	31	25	8– 58
All color defectives	74	15	13	0– 58
Low normals	32	94	100	75–100

TABLE III. Distribution of performance scores. Percentage of each type of color defective subject and of the low normals having different ranges of performance score on Ishihara test (fifth edition).

Type of color defective subject	Number of cases	Percentage of each group achieving scores of Number of plates passed: 0	1–2	3–5	6–8	9–11	12
		%	%	%	%	%	%
Anomalous trichromat							
deuteranomalous	23	30	35	22	13	0	0
protanomalous	12	25	50	25	0	0	0
Dichromat							
deuteranope	15	93	7	0	0	0	0
protanope	17	18	76	6	0	0	0
Low discrimination	7	0	29	43	29	0	0
All color defectives	74	36	41	14	9	0	0
Low normals*	32	0	0	0	0	47	53

* In this group, 53 percent passed all 12 plates (scored 100), 22 percent passed 11 plates (scored 92), 22 percent passed 10 plates (scored 83), and 3 percent passed 9 plates (scored 75).

TABLE IV. Differential classification of deuteranopia and protanopia. Analysis of the value of Plates 12 and 13 (*Series 6*) as a means of differentiating between deuteranopic and protanopic types of red-green defective subjects.

Type of color defective subject	Number of cases	Percentage of cases correctly diagnosed by Plate 12	Plate 13	Both plates	Either plate	Percentage of cases incorrectly diagnosed by Plate 12	Plate 13	Both plates	Either plate	Percentage of cases not diagnosed by Plate 12	Plate 13	Both plates	Either plate
Anomalous trichromat		%	%	%	%	%	%	%	%	%	%	%	%
deuteranomalous	23	65	65	57	74	0	0	0	0	35	35	43	26
protanomalous	12	50	83	50	83	8	0	0	8	42	17	50	8
Dichromat													
deuteranope	15	93	93	87	100	0	0	0	0	7	7	13	0
protanope	17	41	59	29	71	6	6	6	6	53	35	65	23

The overlapping of scores between the various types of defective color vision is brought out still more clearly in Table III which shows the distribution of performance scores for each type. For representation in this table the following classes of performance scores were made: 0 (failure on all plates); 8–17 (passing 1 or 2 plates); 25–42 (passing 3–5 plates); 50–67 (passing 6–8 plates); 75–92 (passing 9–11 plates); and 100 (passing all plates). The overlapping of scores between the various types of color defectives is clearly shown in this table, and consequently the inability of the Ishihara test to yield a classification of either type or extent of defect on the basis of performance score.

The following points are brought out in Tables II and III:

1. The scores of deuteranomalous, protanomalous and low discrimination observers have a wider range of scatter than those of deuteranopes and protanopes. That is, as seems probable, their defect may vary from a slight to an extreme anomaly or low discrimination, while dichromats having a red-green defect would present a more homogeneous group, at least when studied by means of tests designed primarily to detect red-green defect.

2. All color defectives of the red-green types, also cases having low color discrimination amounting to a color defect, are screened from those having normal and low normal color vision if a score of 60 is taken as the critical score on the Ishihara test (fifth edition). In the group tested, no color defective achieved a score of more than 58, and no low normal less than 75.

3. No clean-cut separation as to extent or type of defect can be based on performance scores. While in general the dichromats had lower scores than the anomalous trichromats, a majority of the latter group (65 percent of deuteranomalous and 75 percent of protanomalous) had scores between 0 and 17, the range which includes 100 percent of the deuteranopes and 94 percent of the protanopes tested. *Scores cannot, therefore, be used to indicate either the extent or the type of the color defect.*

4. Of the deuteranopes tested, 93 percent failed all the plates in the Ishihara test; of the protanopes, only 18 percent failed them all. The 5th edition of this test, then, "traps" deuteranopes more decisively than the other types of red-green defect.

3. The Differential Classification of Deuteranopia and Protanopia

So far we have discussed performance scores on the Ishihara test both as a means of detecting the presence of a color defect and as a means of classifying the type of defect, and have shown that the score achieved is adequate to separate color defective from color normal vision when the test is properly administered, but it is not adequate to make the equally important differential classification of type of defect and extent of defect. It remains to discuss *type of response* to certain plates as a clue to differential classification.

In the Ishihara test the plates of *Series 6* are the only ones designed for this purpose. In these plates a double digit is presented ("26" on Plate 12 and "42" on Plate 13). In each plate the left-hand digit is composed of red disks, the right-hand digit of red-purple disks and the background of gray disks. According to Ishihara, the deuteranope type of red-green defect will be able to read only the red digit and the protanope type only the red-purple digit. Our data have been analyzed to determine the effectiveness of these plates for this differential classification.

Difficulty is encountered when other types of response are given, such as the correct reading of both digits or of neither (in which cases of course no diagnosis is possible), or the incorrect reading of one or both digits. In scoring the responses to these plates, the following plan was adopted: A case was called "diagnosed" as to type of defect by the plate in question when one digit was read correctly, and the other not seen at all or read incorrectly. When both digits were read correctly, when neither was read correctly, or when one was read incorrectly and the other not seen at all, the case was called "not diagnosed." In short, "a hit and a miss" are necessary in order to utilize this plate as a test for differential classification. The results obtained are shown in Table IV.

The following points are brought out in Table IV:

1. The plates of *Series 6* are better designed to classify deuteranopes as to type of red-green defect than deuteranomalous observers. The failure in case of the latter group is due either to an incorrect reading of the significant digit or to a correct response to both digits.

2. The plates of *Series 6* are better designed to classify protanomalous observers as to type of color defect than protanopes. The failure in the latter group is due either to an incorrect reading of the significant digit or to inability to see any digit.

3. Plates 12 and 13 give equally good results for the deuteranopic type, but Plate 13 is better than 12 for distinguishing the protanopic type. (83 percent and 59 percent with Plate 13 as compared with 50 percent and 41 percent for Plate 12.)

4. One protanomalous case and one protanope of our group would have been classified as deuteranope by one or both plates of *Series 6;* under Mazda light this number would have been greatly increased. None of the deuteranopic type was incorrectly classified.

5. It is seen, for example, that whereas 41 percent of the protanopes were correctly typed (a hit and a miss) by Plate 12, 59 percent were correctly typed (a hit and a miss) by Plate 13. However, if the indicated response were required from each plate (i.e., a hit and a miss on each plate), only 29 percent of the protanopes were correctly typed. On the other hand, if the indicated response were required from only one of the two plates, ignoring the results obtained from the other, then 71 percent of the protanopes were correctly typed—the second plate yielding an ambiguous response.

CONCLUSIONS

The following conclusions seem justified concerning the fifth edition of the Ishihara test:

1. Properly administered the test affords a good rough device for screening red-green defectives from color normal individuals if a performance score of 60 is taken as the critical score. In the group tested no color defective scored more than 58 and no low normal less than 75.

2. No analysis as to type or extent of defect can be based on performance scores.

3. The test yields a lower score by the deuteranopic than the protanopic type of color defective. (This is, in effect, a "weighting" which is lost under more yellowish illumination.)

4. The plates of *Series 6*, which were designed to distinguish between deuteranopes and protanopes, are adequate to classify deuteranopes alone, when the report on either of the two plates is taken as indicative. On this basis only 74 percent of the deuteranomalous cases are correctly typed, 83 percent of the protanomalous cases and 71 percent of the protanopes.

5. The Ishihara test is a gross test for defective red-green vision. It fails to classify the type of defective color vision (protanope, protanomalous, deuteranope, deuteranomalous, tritanope, tritanomalous) and cannot be used to give a satisfactory evaluation of the extent or degree of defect, no matter how carefully administered.

6. Plate 4 is a practically perfect screening test when properly used, but if used with Mazda light its efficiency drops from 100 percent to 43 percent for deuteranomalous cases.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 4 APRIL, 1945

Thickness Control of Sharp Cut-Off Type of Glasses*

H. P. GAGE
Corning Glass Works, Corning, New York

It has been noted that certain spectrophotometric curves, especially those of the sharp cut-off type of colored glasses, have essentially the same shape—merely displaced in wave-length for different thicknesses. If this were strictly true, it must be that the curve between log log transmittance and wave-length is a straight line, since only a straight line can be moved vertically (by change of glass thickness) and match the same curve moved horizontally (by change of wave-length). If it is assumed that the log log transmittance *versus* wave-length curve is a straight line the curve between transmittance and wave-length has an S shape which is not a bad fit for actual spectrophotometric curves of certain glasses. If the fit were perfect, the wave-length displacement of the curve caused by change in thickness would be proportional to the change of the logarithm of thickness, the constant in the equation can be determined by the slope of the curve. With a series of infinitely sharp cut-off curves in which there is complete transmission for all wave-lengths greater than a certain wave-length, and complete absorption for all shorter than this, the values of Y, x, y, and z (Transmission and x, y, and z I.C.I. Coordinates), may be calculated. A wave-length of cut can be found which will produce the same chromaticity, x and y coordinates as any of the sharp cut glasses but the transmission will differ, being the maximum possible for the given chromaticity. The change in chromaticity expressed as change in y coordinate can be determined relative to change in wave-length of cut-off. By combining this latter relation with the change of wave-length of cut-off caused by change of thickness of the given glass as determined from the shape of the log log curve it is possible to calculate the required change in thickness to produce any desired change in the y coordinate within the capabilities of the glass.

IT has been noted in the examination of a large number of spectrophotometric curves of glasses where several different thicknesses of glass were plotted on the same sheet that with certain types of glass the shape of all the curves are essentially the same.

With increasing wave-length the transmission increases abruptly with but a slight "toe" along a straight line, then through a moderate shoulder to a maximum. This shape occurs for a long series of "sharp cut-off" yellows, oranges and reds (Fig. 1). As the thickness of one of these glasses is reduced, there results a change of color and transmission much less noticeable than occurs with most other kinds of colored glass; but a small change does occur and, for obtaining standards or other pieces to meet exact requirements, it would be convenient to know how much these properties can be controlled by thickness. A hint as to a possible means of attack is to be found in A. C. Hardy's *Handbook of Colorimetry*[1] where it is pointed out that if the spectrophotometric curve is plotted with log log transmittance for abscissae the curve will always have the same shape, being merely displaced up or down by change of thickness. If now, on the ordinary type of plot of transmission *versus* wave-length, the change of thickness results in a bodily shift of the curve in wave-length but not in shape, then it must be that the log log of internal transmittance must be a straight line, as only a straight line can be moved vertically as by a change in glass thickness and coincide with the same curve moved horizontally as by a change in wave-length. The course of the abbreviated mathematical analysis to follow will be first to determine the mathematical properties of a curve which satisfies the requirement that log log T' *versus* λ is a straight line, then to determine the shape of the corresponding curve between T' and λ, and then to determine whether the resulting S shape curve corresponds to that of actual measured glasses. For those glasses whose actual transmission curves fit the theoretical curves most perfectly, any mathematical deductions from the theoretical curve will apply to the actual glasses, and where the data cannot be made to fit the curve, there will be departures from an exact correspondence with those deductions.

I. SYMBOLS USED

T is the transmission of the glass at a given wave-length.

* Paper presented at the Twenty-Ninth Annual Meeting of the Optical Society of America, New York, New York, October 20 and 21, 1944.

[1] A. C. Hardy, *Handbook of Colorimetry* (The Technology Press, Massachusetts Institute of Technology, Cambridge Massachusetts, 1936).

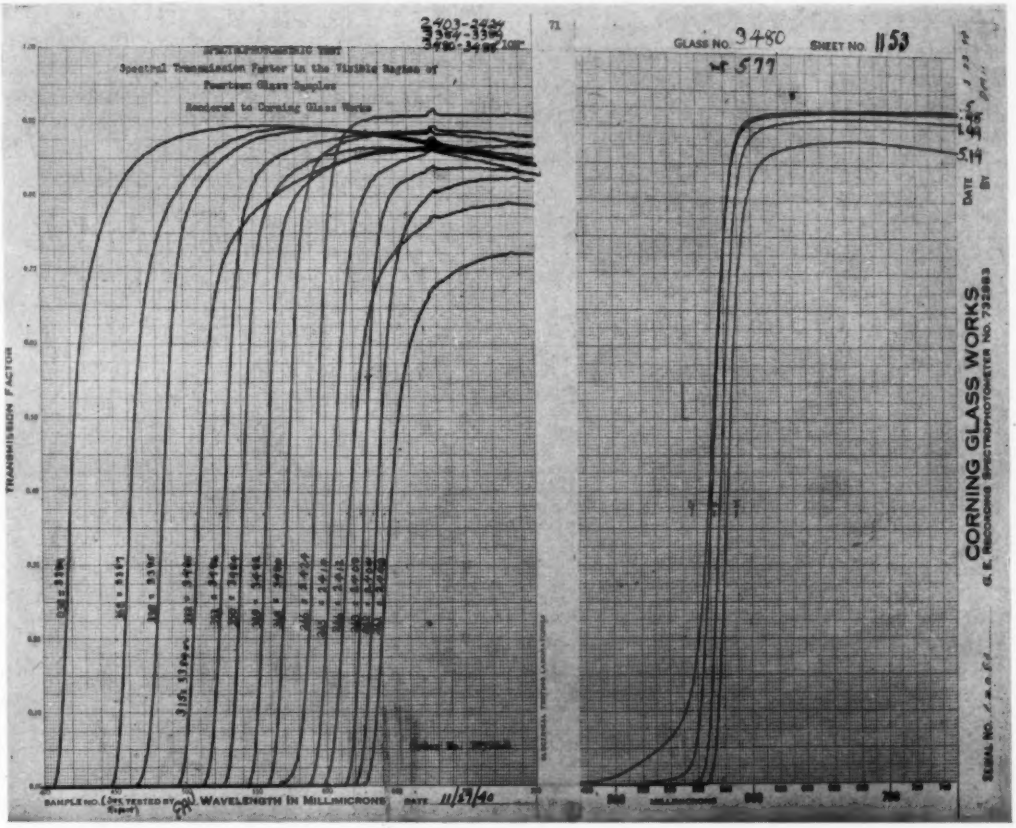


FIG. 1. (a) Sharp cut-off type glasses, reproduction of curve sheet made on Hardy G.E. spectrophotometer from nearly colorless Noviol A no. 3389 to dark Pyrometer red no. 2403. The two glasses with lowest transmission at long wave-length end are of the same melt but different thickness. All other glasses are of separate melts. (b) Four pieces of the same melt of an orange glass, thickness roughly .5 mm, 1 mm, 2 mm, and 5.14 mm. The .5 mm piece is too thin to have complete and abrupt absorption at the short wave-length end and has a gradual toe. The others show the small but definite shift in cut-off with increasing thickness. Also, that the curves rise to a nearly flat maximum with but small absorption for the longer wave-lengths.

h is T_{max}, the maximum transmission at any wave-length.

T' is T/T_{max}.

T'' is the internal transmittance of the glass, i.e., corrected for surface reflection, which is about $T/0.92$.

t is thickness.

λ is wave-length.

βt is $\log_{10} T''$ or $\log T'$ which ever applies.

γ is $\log_{10} \beta t = \log_{10} \log_{10} T'$.

n is $\Delta y/\Delta\lambda$ from Table II.

$\Delta\lambda$, $\Delta\gamma$, etc. are used because differences in these factors can be appreciable when dealing with straight lines or curves from the applicable data in Table II.

Properties of transmission curves when γ *versus* λ is a straight line.

In Table I and Fig. 2 are shown the values of T' and βt for various values of $\gamma = \log \log T'$.

Temporarily disregarding signs and comparing the numerical values, it is noted that points in Column 4 calculated as $.8\gamma + .08$ are near the corresponding values of T' between the limits of $T' = .15$ and $.63$ and that great departures occur for larger and smaller values of T'. The straight line having the equation $y = .8\gamma + .08$

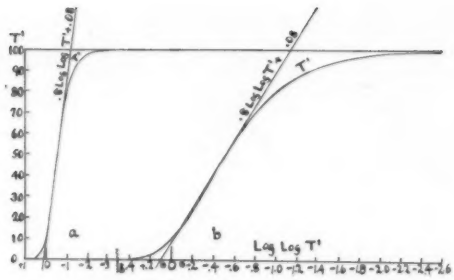


FIG. 2. A mathematical curve illustrating the relationships in Table I. On the horizontal axis is plotted the values of log log T' and along the vertical axis is plotted the corresponding values of T' 0 to 1.00. Also, a straight line representing .8 log log T'+.08 which it is noticed nearly coincides with the S shape T' curve between the limits of T'=.15 and .64 but the T' curve departs from the straight line at the toe below .16 and on the long shoulder above .64. The differences are to be noted in the last two columns of Table I. Curve a is condensed in horizontal scale and has a close resemblance to the actual spectrophotometric curves in Fig. 1; curve b is on a more extended horizontal scale.

approximates the S shape curve $y=T'$ between the above-mentioned limits.

The relation .8γ+.08 approximates T' can be written γ approximates 1.25T'−.1 between limits γ=.1 to γ=.7 or T'=.15 to T'=.63.

Certain other relations between the values of T', βt and γ may be calculated but are not used in the argument here presented but may be noted below.

TABLE I.

γ = log log T'	βt = log T'	T'	Log log T' ×.8 +.08
⊙+ .5	−3.162	.0007	
⊙+ .4	−2.512	.0031	
⊙+ .3	−1.995	.0101	
⊙+ .2	−1.585	.026	
⊙+ .1	−1.259	.055	.00
⊙ .00	−1.00	.100	.08
⊙− .1	− .7943	.1606	.16
⊙− .2	− .6310	.2339	.24
⊙− .3	− .5012	.3154	.32
⊙− .4	− .3981	.3998	.40
⊙− .5	− .3162	.4829	.48
⊙− .6	− .2512	.5607	.56
⊙− .7	− .1995	.6317	.64
⊙− .8	− .1585	.6942	.72
⊙− .9	− .1259	.7485	.80
⊙−1.0	− .1000	.7943	.88
⊙−1.1	− .0794	.8325	.96
⊙−1.2	− .0631	.8650	1.04
⊙−1.3	− .0501	.8910	
⊙−1.5	− .0316	.930	
⊙−1.7	− .0199	.955	
⊙−2.0	− .0100	.9772	
⊙−2.3	− .0050	.9886	
⊙−2.5	− .0032	.9930	
⊙−2.7	− .002	.9954	
⊙−3.0	− .001	.9977	
⊙− .5215	− .3010	.500	.498
⊙− .3622	− .4343	.3678	.370
⊙− .3010	− .500	.3162	.321
⊙− .1555	− .699	.200	.2045
⊙+ .1142	−1.3010	.05	
⊙+ .2302	−1.699	.02	

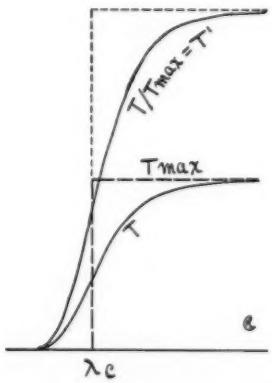


FIG. 3. Measured spectral transmission curve and $T'=T/T_{max}$ curve. The derived or T' curve has the same relative transmission distribution as the measured transmission curve but for each wave-length the values are increased sufficiently so the maximum point is 100 percent or unity. The chromaticity of light transmitted with either, i.e., the x, y, and z coordinates, will be the same and also the same as a theoretical transmitter in which all wave-lengths shorter than the cut-off wave-length λc are removed, and all wave-lengths greater than λc are completely transmitted. The luminous transmissions, i.e., the Y values, will differ.

Note. The transmission of a substance is always less than unity. The logarithm of a number less than unity is negative. Thus log 2 is +.301030 while log .2, usually expressed as −1+.301030 or 9.30103–10 is −.698970 or log 1/5. In many discussions this is referred to as the logarithm of the reciprocal but I prefer to carry out the mathematics in the positive form and later note whether the result is positive or negative. Similarly log .02 (log 2 percent) is −1.698970. For conveniently determining the logarithms of numbers less than unity the equal division scale of a slide rule may be inserted in the reverse direction or a four place table of antilogarithms may be renumbered in the reverse direction.

The log log T (logarithms of logarithms) of numbers less than unity involve logarithms of negative numbers. Thus log 0.2 is −0.698970, log log 0.2 is log −0.698970, i.e., ⊙−0.15554. The circle dot sign ⊙ indicates that the logarithm of a negative number is involved and the minus sign "−" indicates that the logarithm applies to a number less than unity.

In preparing the working tables here included (Table I) the simplest direction to work with was to assume equally spaced values of log log T' and later determine the corresponding values of log T' and T'.

Derivative Forms

Natural logarithms, to the base e, are convenient for mathematical transformation but on account of the tables being clumsy to use and frequently being unavailable, the same transformations can be handled using the more usual type to the base 10. Thus if

$$10^x = a; \qquad 10^{\beta t} = T'. \tag{1}$$

$$X = \log_{10} a; \qquad \beta t = \log_{10} T'. \tag{2}$$

$$d\,10^x = 10^x \log_e 10\, dx; \quad d\,10^{\beta t} = 10^{\beta t} \log_e 10\, d\beta t; \tag{3}$$

$$= 10^x\, dx/.4343; \quad d\,10^{\beta t} = 10^{\beta t} d\beta t/.4343.$$

i.e.,

$$dT'/d\beta t = T'/.4343. \tag{4}$$

Similarly let

$$\gamma = \log_{10} \beta t, \tag{5}$$

i.e.,

$$10^\gamma = \beta t, \tag{6}$$

and

$$d\beta t/d\gamma = \beta t/.4343. \tag{7}$$

As

$$\frac{dT'}{d\beta t} \times \frac{d\beta t}{d\gamma} = \frac{dT'}{d\gamma}, \tag{8}$$

we get

$$\frac{T'}{.4343} \times \frac{\beta t}{.4343} = \frac{dT'}{d\gamma} = \frac{T'\beta t}{.4343^2} = \frac{T'\beta t}{.18861}. \tag{9}$$

Second derivative $d^2T'/d\gamma^2$.

From (9) write

$$dT' = T'\beta t\, d\gamma/.4343^2,$$

and (7)

$$d\beta t = \beta t\, d\gamma/.4343$$

and,

$$d\,UV = U\,dv + V\,du.$$

$$d^2T' = d(T'\beta t\, d\gamma)/.4343^2, \tag{10}$$

$$= (dT'\beta t\, d\gamma + T'\, d\beta t\, d\gamma)/.4343^2. \tag{11}$$

Substitute value of dT' from (9) and $d\beta t$ from (7) in (11) we get

$$d^2T' = \frac{1}{.4343^2} \times \left(\frac{T'\beta t\, d\gamma \times \beta t\, d\gamma}{.4343^2} + \frac{T'\beta t\, d\gamma\, d\gamma}{.4343} \right), \tag{12}$$

i.e.,

$$d^2T'/d\gamma^2 = \frac{T'}{.4343^3} \left(\frac{\beta t}{.4343} + 1 \right). \tag{13}$$

Equation (13) is important in showing that the curve for $d^2T'/d\gamma^2$ has a zero value when $\beta t/.4343 + 1 = 0$, i.e., $\beta t = -.4343$, hence the curve for $dT'/d\gamma = dT'/d \log \log T'$ has a maximum at this point and the steepest part of the curve of T' plotted against $\log \log T'$ occurs at this point, i.e., where $\log \log T' = \odot - .3622$, $\log T' = -.4343$, $T' = .3678$, and also at this point the slope of the curve is according to Eq. (9)

$$dT'/d\gamma = T'\beta t/.4343^2 = (T' \times -.4343)/.4343^2$$
$$= T'/-.4343 = -.848.$$

The rough approximation for the nearly straight portion of the curve calls for a slope of .800.

II. FINDING THE HYPOTHETICAL CURVE APPROXIMATING AN ACTUAL CURVE

The actual transmission T curves of the sharp cut-off glasses cannot follow the equation that $\log \log T$ *versus* λ is a straight line as this would involve a 100 percent transmission for very long wave-lengths. This cannot occur on account of (1) surface reflection and (2) non-selective absorption in the body of the glass often referred to as "mud." These factors may be compensated by using corrected values. They are referred to as $T' = T/T_{max}$ where T_{max} is the maximum transmission for any wave-length. The T' curve is therefore a shape factor.

It is to be noted that two spectral curves (Fig. 3) having all transmissions in the same ratio will give the same chromaticity (x, y, and z values), although they will correspond to different transmissions or Y values. The values of T_{max} will of course follow the thickness relation that $\log T_{max}'' = \beta_{max}\, t$ where T_{max}'' is T_{max} corrected for surface reflection only.

The first step is to convert a sufficient number of points on the measured transmission curve to the T' values, thence determine the γ values, or, more conveniently, to take T' values corresponding to chosen γ values from Table I, determine the reduced T values ($T = T' \times T_{max}$), and note at what wave-length these occur. The essential information is to be obtained from the straight line which is followed by the points $\log \log T'$ plotted against wave-length, and to define this line but two points need be determined.

In case curves are available for two or more

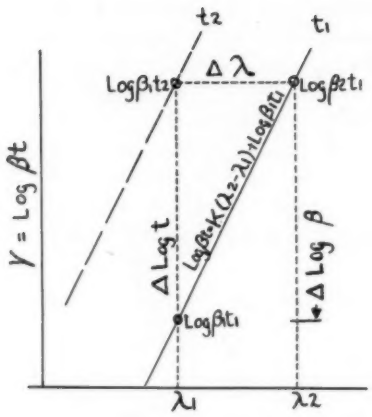


FIG. 4. The log log T' is plotted against wave-length λ for two pieces of the same glass of differing thickness, assuming these curves to be straight lines. For a single glass the difference in height of the points is $\Delta \log \beta$. For two thicknesses of glass having the same transmission for two different wave-lengths, λ_1 and λ_2 differing by $\Delta\lambda$ the vertical separation is $\Delta \log t$.

thicknesses of glass it is theoretically not enough to determine merely the variation of wave-length at which each curve passes through a given transmission, as say 37 percent, but to find the maximum transmission for each, and determine the wave-lengths for which the transmission is some fraction, say 40 percent, of the maximum.

The straight line or γ curves are for purposes of this study most conveniently expressed as log βt curves *versus* wave-length (Fig. 4) where $\beta t = \log T'$.

With the two points on a single curve for thickness t_1 known, k, the slope can be determined as

$$k = (\Delta \log \beta t)/\Delta\lambda,$$

or

$$k = (\log \beta_2 t_1 - \log \beta_1 t_1)/(\lambda_2 - \lambda_1);$$

but as $\log \beta t = \log \beta + \log t$, we get

$$k = (\log \beta_2 + \log t_1 - \log \beta_1 - \log t_1)/(\lambda_2 - \lambda_1).$$

Eliminating log t, we get

$$k = (\log \beta_2 - \log \beta_1)/(\lambda_2 - \lambda_1), \quad k = \Delta\beta/\Delta\lambda.$$

If curves for two or more thicknesses are available the wave-length λ_1 for the second thickness t_2 which has the same value of T' as does the first thickness t_1 at λ_2 is determined. The value of T' for this same wave-length λ_1 is assumed as known although it is eliminated from the equations.

The new point $\log \beta_1 t_2$ corresponds to wave-length λ_1 but is equal to $\log \beta_2 t_1$. This when substituted in the equation for the slope of both γ lines gives

$$k = \frac{\log \beta_2 t_1 - \log \beta_1 t_1}{\lambda_2 - \lambda_1} = \frac{\log \beta_1 t_2 - \log \beta_1 t_1}{\lambda_2 - \lambda_1},$$

but $\log \beta t = \log \beta + \log t$; hence,

$$k = \frac{\log \beta_1 + \log t_2 - \log \beta_1 - \log t_1}{\lambda_2 - \lambda_1};$$

eliminating β_1 which is the same for both thicknesses

$$k = \frac{\log t_2 - \log t_1}{\lambda_2 - \lambda_1} \quad \text{or} \quad k = \frac{\Delta \log t}{\Delta\lambda}.$$

It is thus seen that equal shifts in wave-length of the cut-off are caused by equal changes in log thickness.

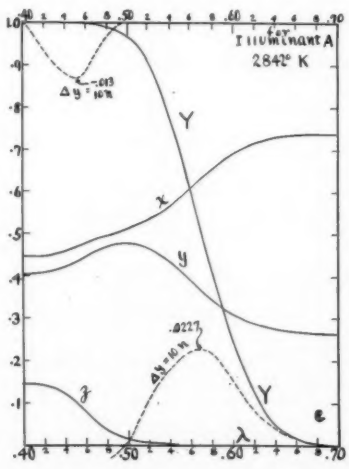


FIG. 5. Relationships shown in Table II. For theoretical sharp cut-off filters having zero transmission at the wave-length indicated and complete transmissions for all greater wave-lengths the trichromatic coefficients x, y, and z and the transmission Y are as shown. This is calculated for illuminant A, 2842°K. The dashed curve Δy is the difference in the y value caused by a change in cut-off position by 0.01 micron or 10 mμ. The slope n, i.e., $\Delta y/\Delta\lambda$ when λ is measured in milimicrons is 1/10 of this value.

TABLE II. Colorimetric properties of theoretical sharp cut filters with illuminant A, at 2842°K.

λ	$\overleftarrow{\bar{x}E}$	$\overrightarrow{X}$	$\overrightarrow{x}$	$\overleftarrow{\bar{y}E}$	$\overrightarrow{Y}$	$\overrightarrow{y}$	$\overleftarrow{\Delta y=10n}$	$\overleftarrow{\bar{z}E}$	$\overrightarrow{Z}$	$\overrightarrow{z}$	$\overrightarrow{\Sigma}$	λ'
770	2	2		0								765
760	4	6	.75	2	.00002						8	755
750	6	12	.75	2	.00004						16	745
740	14	26	.725	6	.00010						36	735
730	28	54	.73	10	.00020						74	725
720	57	111	.740	19	.00039	.26000					150	715
710	110	221	.7367	40	.00079	.26333					300	705
700	209	430	.7363	75	.00154	.26370	.00037				584	695
690	404	834	.7354	146	.00300	.26455	.00085				1134	685
680	804	1638	.7345	292	.00592	.26547	.00092				2230	675
670	1448	3086	.7334	530	.01122	.26663	.00116				4208	665
660	2628	5714	.7318	972	.02094	.26879	.00156				7808	655
650	4336	10050	.7293	1637	.03731	.27074	.00255				13781	645
640	6558	16608	.7252	2562	.06293	.27479	.00405				22901	635
630	8980	25588	.7191	3704	.09997	.28093	.00614				35585	625
620	11373	36961	.7103	5071	.15068	.28959	.00866	3	3	.00006	52032	615
610	12669	48630	.6984	6356	.21424	.30149	.01190	4	7	.0001	71061	605
600	12704	62334	.6826	7547	.28971	.31724	.01575	10	17	.0002	91322	595
590	11579	73913	.6631	8540	.37511	.33656	.01932	12	29	.0003	111453	585
580	9719	83632	.6412	9228	.46739	.35838	.02182	18	47	.0004	130418	575
570	7571	91203	.6185	9457	.56196	.38104	.02266	21	68	.0005	147467	565
560	5510	96713	.5961	9222	.65418	.40323	.02219	36	104	.0006	162235	555
550	3732	100445	.5753	8568	.73986	.42372	.02049	75	179	.0010	124610	545
540	2313	102758	.5564	7600	.81586	.44176	.01804	162	341	.0018	184685	535
530	1214	103972	.5400	6322	.87908	.45659	.01483	309	650	.0034	192530	525
520	425	104397	.5226	4771	.92679	.46748	.01089	525	1175	.0059	198251	515
510	57	104454	.5162	3080	.95759	.46322	+.00574	969	2144	.0106	202357	505
500	27	104481	.5080	1792	.97551	.47427	+.00105	1509	3653	.0177	205685	495
490	160	104641	.5002	1039	.98590	.47125	−.00302	2324	5977	.0285	209208	485
480	428	105069	.4912	622	.99212	.46384	−.00741	3636	9613	.0450	213894	475
470	776	105845	.4808	362	.99574	.45230	−.01154	5116	14729	.0669	220148	465
460	1019	106864	.4703	210	.99784	.43914	−.01316	5851	20580	.0906	227228	455
450	1031	107895	.4615	117	.99901	.42727	−.01187	5435	26015	.1113	233811	445
440	926	108821	.4545	61	.99962	.41747	−.00980	4647	30662	.1280	239445	435
430	649	109470	.4500	27	.99989	.41099	−.00648	3167	33829	.1390	243288	425
420	262	109732	.4482	8	.99997	.40846	−.00253	1256	35085	.1433	244814	415
410	71	109803	.4478	2	.99999	.40778	−.00068	340	35425	.1445	245227	405
400	19	109822	.4476	1	1.00000	.40760	−.00018	93	35518	.1448	245340	395
390	5	109827	.4476	0				23	35541	.1448	245368	385
380	1	109828	.4476	0				6	35547	.1449	245375	375

III. VARIATION OF CHROMATICITY WITH CUT-OFF

The variation of chromaticity with the wave-length position of the cut-off can be approximated by the method described by the author before the Optical Society of America.[2] The results described at the time applied to the O.S.A. colorimetric data[3] and were never published in full as the O.S.A. trichromatic coefficients were so soon superseded by the more comprehensive data incorporated in those of the I.C.I. standard observer. The same method of attack, however, is valid. A similar method was used by MacAdam,[4] but was not carried out fully for colors in which it was assumed that for all wave-lengths greater than the cut-off point there is complete transmission, and for all shorter wave-lengths there is zero transmission. With such a simple assumption the calculation of transmission and trichromatic coefficients reduces to an easy routine. These are shown in Table II.

The headings of the columns are: λ, wave-length. This is the wave-length for which the values of $\bar{x}E$, $\bar{y}E$, and $\bar{z}E$ apply in the tables for calculating the I.C.I. coordinates of light transmitted by filters. They represent the middle of a 10 mμ band of added light.

[2] H. P. Gage, "Colorimetric Efficiency of Red Glass," J. Opt. Soc. Am. **21**, 433 (1931).

[3] Report of the Committee on Colorimetry for 1920–21, J. Opt. Soc. Am. **6**, 527 (1922).

[4] D. L. MacAdam, "The Theory of the Maximum Visual Efficiency of Colored Materials," J. Opt. Soc. Am. **25**, 249 (1935); and **25**, 361 (1935).

Columns $\bar{x}E$, $\bar{y}E$, and $\bar{z}E$ are the three sensitivities of the eye for the three factors $\bar{x}$, $\bar{y}$, and $\bar{z}$ multiplied by the relative energy of the light source, in this case Illuminant A, at 2842°K, for the same band of wave-lengths. These are copied from the tables appearing in Judd's paper[5] and also the report of the Optical Society Committee on Colorimetry.[6]

X, Y, and Z are the running sums of all numbers in the $\bar{x}E$, $\bar{y}E$, and $\bar{z}E$ columns as read from an adding machine. These running sums constitute the X, Y, and Z values of the hypothetical filter having complete transmission for the longer wave-lengths and complete absorption for the shorter wave-lengths. The values in Column Y represent the luminous transmission of such filters.

$\sum$. The column headed $\sum$ is the horizontal sum X, Y, and Z.

x, y, and z are $X/(X+Y+Z)$, $Y/(X+Y+Z)$ and $Z/(X+Y+Z)$ and are the I.C.I. coordinates of the color in question.

$\Delta y = 10n$. With the sharp cut-off type of absorption the chromaticity is rather uniquely specified by the y value so this column is more completely studied. The difference in y factor with wave-length ($10\Delta y/\Delta\lambda$) is calculated by subtracting successive y values which gives the average slope of the curve, but in this case it is the average slope for which the original wave-length column applies, thus adding the 575–565 band changes y from .35838 to .38104, i.e., .02266, but this divided by 10 is the average slope $\Delta y/\Delta\lambda$ at the 10 mμ band centering at 570. This factor $\Delta y/\Delta\lambda$ is called n for convenience.

λ'. It is necessary to use care in determining the position of cut-off to which these coordinates apply. Thus consider the effect of adding the ($\bar{x}E$) values for say λ570 to the previous values, we are really adding the effect of a band of energy extending from λ575 to λ565 so the resulting color corresponds to a sharp cut-off at the shorter wave-length, 5 mμ shorter than the wave-length opposite the last figures added, i.e., we are adding a band which extends the effect of the previous cut-off at λ575 to include all energy to a sharp cut-off at λ565. With this in mind is added a new column of wave-lengths to indicate correctly the position of the cut-off. The arrows at the head of the columns indicate whether to use the λ or λ' wave-length.

IV. THICKNESS AND CHROMATICITY

From Part II we know the relation between change in thickness and change in the position of the cut-off; from Part III we can determine the relation between position of cut-off and chromaticity, therefore the relation between change in thickness and change in chromaticity can be determined.

$$\Delta\lambda/\Delta \log t = 1/k.$$

$n = \Delta y/\Delta\lambda$ from Table II.

$$\Delta y/\Delta\lambda \times \Delta\lambda/\Delta \log t = \Delta y/\Delta \log t,$$

or

$$\Delta y/\Delta \log t = n/k \quad \text{or} \quad \Delta y = \frac{n}{k} \Delta \log t.$$

[5] Deane B. Judd, "The 1931 Standard Observer and Coordinate System for Colorimetry," J. Opt. Soc. Am. **23**, 359–374 (1933). See Table III, p. 365.

[6] Committee on Colorimetry, "Quantitative Data and Methods for Colorimetry," J. Opt. Soc. Am. **34**, 633–688 (1944). See Table VII, p. 643 .

The Cos⁴ Law of Illumination*

MAX REISS
Eastman Kodak Company, Hawk-Eye Works, Rochester, New York

IT is well known that for the design of wide-angle lenses it is of importance to know the light distribution over the image plane. In spite of the fact that the relationship between illumination and field angle is derivable from simple principles, this relationship, commonly referred to as the cos⁴ law, is often found to be incorrectly interpreted in the optical literature. This fact has been pointed out in an interesting paper by G. Slussareff.[1] He emphasizes the importance of considering the effect of the aberration of the pupils on the illumination, and he derives an expression based on Seidel theory for the light distribution produced by a system with an internal diaphragm. He concludes that it might be possible to construct a system in such a way that the illumination will be much better than according to the cos⁴ law.

The purpose of the present paper is, first, to draw attention again to some of the paradoxical results which are obtained if the theoretical laws are misinterpreted, and furthermore, to show that the relationship between the entrance and exit pupils of a system derived by considerations of the flow of energy, also follows from purely optical laws (theorems of Clausius and Straubel). It will also be shown that the assumptions made by Slussareff for the derivation of his law of illumination seem to be such as to deprive it of practical value.

THE LAMBERT AND STRAUBEL THEOREMS

The light energy (flux) radiated per second from a small surface element df_1 to an element df_2 (Fig. 1) is given by

$$d\Phi = B \cdot df_1 \cos\theta_1 d\Omega_1 \quad \text{(Lambert's law)}, \quad (1)$$

FIG. 1.

where B is the intrinsic brightness of df_1, $d\Omega_1$ is the solid angle subtended by df_2 at df_1, and θ_1 is the angle between the normal to the surface element df_1 and the line connecting the surface elements.

Let us now consider a lens system, df and df' being small surface elements of the object and image planes (Fig. 2), and let us assume that the points on df are imaged in corresponding points of df'. The flux from df carried in a thin pencil is

$$d\Phi = B \cdot df \cos W \cdot d\Omega. \quad (2)$$

We exclude losses by reflection and absorption inside the system. The emerging pencil going towards df' contains the flux

$$d\Phi' = B' df' \cos W' d\Omega', \quad (3)$$

where B' and $d\Omega'$ correspond to the values B and $d\Omega$ in the object space. The entering flux must be equal to the emerging flux, hence

$$B \cdot df \cdot \cos W \cdot d\Omega = B' df' \cos W' d\Omega'. \quad (4)$$

Let us now draw attention to a theorem stated by Straubel[2] which can be derived from Fermat's principle. Consider a thin pencil of light (Fig. 3) emanating from the point A (solid angle $d\Omega$) and then emerging from an optical system and intersected by any plane E_2. The area bounded by the intersection between the pencil and the plane E_2 is df'. Suppose now that from the point B, which is the intersection point of the axis of the pencil with the plane E_2, there emanates a thin pencil (solid angle $d\Omega'$) along the same axis

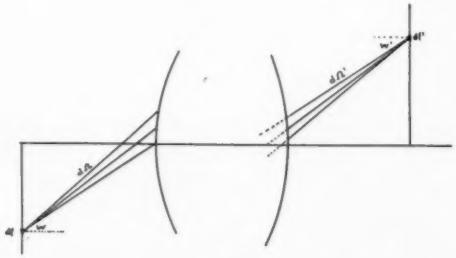


FIG. 2.

* Paper presented at the Twenty-Ninth Annual Meeting of the Optical Society of America, New York, New York, October 20 and 21, 1944.

[1] G. Slussareff, J. Phys. U.S.S.R. **4**, No. 6, 537 (1941).

[2] R. Straubel, Physik. Zeits. **4**, 114 (1902–03).

as before, going backwards through the system. This pencil cuts out an area df from any plane E_1 containing A. If θ and θ' are the respective angles of the axis of the pencil with the normals to planes E_1 and E_2, then Straubel's theorem states

$$n^2 df \cdot d\Omega \cos\theta = n'^2 df' d\Omega' \cos\theta', \qquad (5)$$

(n is the index of refraction of the first medium, and n' that of the last medium).

In the foregoing discussion we did not assume that the points on df' were the optical images of the points on df. For such a case Straubel's theorem in the above form loses its meaning because A and B would become merely conjugate points.

By using this theorem we can derive another theorem (Fig. 4). Let us first apply Straubel's theorem to a point A on df and a point F on any auxiliary plane E. If the pencil emanating from F is chosen such that it includes df then

$$n^2 df \cdot d\Omega \cos\theta = \bar{n}^2 d\bar{f} \cdot d\bar{\Omega} \cos\bar{\theta}. \qquad (6)$$

Let us now apply the same theorem a second time, namely, between the point F and the point A' (image of A). The pencil emanating from F is the same as the one going from F towards df. Since df' is the image of df this pencil includes df'. The pencil leaving A' is the pencil which includes $d\bar{f}$ and then unites in A. We therefore have

$$\bar{n}^2 d\bar{f} \cos\bar{\theta} \cdot d\bar{\Omega} = n'^2 df' \cos\theta' d\Omega'. \qquad (7)$$

By combining the two equations we get

$$n^2 df \cos\theta d\Omega = n'^2 df' \cos\theta' d\Omega'. \qquad (8)$$

Formula (8) looks exactly like (5) but the meaning is different. Here θ and Ω on one side, and θ' and Ω' on the other side, belong to the same pencil. Equation (8) was derived by Clausius[3] by a different argument.

If we now apply Eq. (8) to Fig. 2, then we have for a system in air

$$df \cdot \cos W \cdot d\Omega = df' \cos W' d\Omega'. \qquad (9)$$

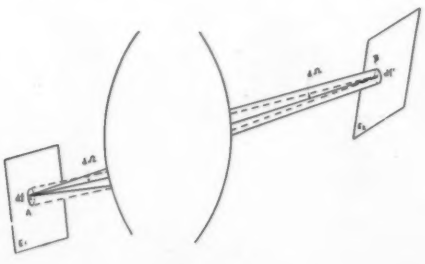


FIG. 3.

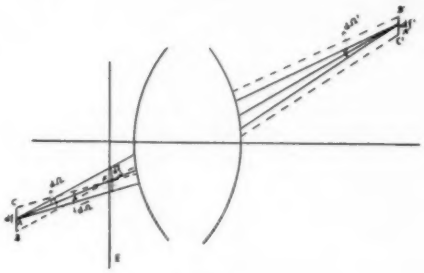


FIG. 4.

By comparing (9) with (4) we have

$$B = B'. \qquad (10)$$

This well-known result is usually derived only for a point on the axis, by use of the sine theorem. The present derivation from the Clausius-Straubel theorem is obviously much more general.

The flux can be written as either

$$B \cdot df \cos W \cdot d\Omega \qquad (11)$$

or

$$B \cdot df' \cos W' d\Omega'. \qquad (12)$$

THE RELATIONSHIP BETWEEN ENTRANCE AND EXIT PUPILS

Let us consider again a small pencil of light radiated from a surface element df (Fig. 5). The cross section of the pencil with any vertical plane E is S. Then

$$d\Omega = \frac{S \cos W}{(e/\cos W)^2}.$$

If the corresponding pencil in the image space intersects another vertical plane E' at S', then

$$d\Omega' = \frac{S' \cos W'}{(e'/\cos W')^2}.$$

Since the values of $d\Omega$ and $d\Omega'$ do not depend upon the situation of E and E' we can place

[3] R. Clausius, *The Mechanical Theory of Heat* (The Macmillan Company, London, 1879), p. 322.

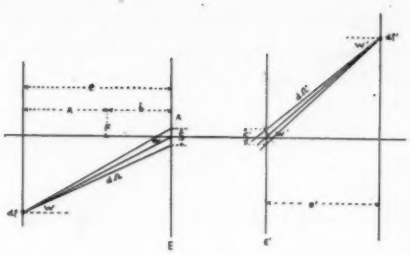


FIG. 5.

them where we please. For example, E may be at the crossing point of the principal ray with the axis in the object space. Then from (9) we have

$$\frac{df\cdot\cos W\cdot S\cdot\cos W}{(e/\cos W)^2}=\frac{df'\cos W'\cdot S'\cos W'}{(e'/\cos W')^2}, \tag{13}$$

or

$$\frac{S'}{S}=\frac{df}{df'}\cdot\frac{\cos^4 W}{\cos^4 W'}\cdot\frac{e'^2}{e^2}. \tag{14}$$

Let us assume that df/df' is constant over the field (distortion free), and let us further assume for the moment that we have a fixed diaphragm in front of the system (S and e constant); then, if C represents a constant:

$$S'=C\left(\frac{\cos W}{\cos W'}\right)^4. \tag{15}$$

This means that with varying obliquity S' generally changes its size (distortion of the exit pupil).

We are especially interested in an object at infinity. For this case (14) has to be transformed since $e=\infty$, and df and df' lose their usual meaning.

For a finite object distance, the true magnification is

$$M=(df'/df)^{\frac{1}{2}}. \tag{16}$$

Furthermore, let

$$m=F/x \quad \text{(Gaussian magnification)}, \tag{17}$$

and

$$m_p=F/b$$

(paraxial magnification of the pupils). (18)

Since

$$e=b-x, \tag{19}$$

$$e=F(1/m_p-1/m)=F\frac{m-m_p}{mm_p}. \tag{20}$$

If the system is free from distortion, then $M=m$, and we have from (14)

$$\frac{S'}{S}=\frac{\cos^4 W}{\cos^4 W'}\frac{e'^2}{F^2}\left(\frac{m_p}{m-m_p}\right)^2. \tag{21}$$

Since for an object at infinity $m=0$, we simply get

$$\frac{S'}{S}=\frac{e'^2}{F^2}\cdot\frac{\cos^4 W}{\cos^4 W'} \text{ (for an object at infinity)}. \tag{22}$$

This is a general relationship between the cross sections S and S' of the entering and emerging pencils with any two planes in the object and image space. For the point on the axis $\cos W=\cos W'=1$, and we get the paraxial magnification formula

$$\frac{S'}{S}=\left(\frac{e'}{F}\right)^2. \tag{}$$

We can now derive the illumination at df'. Since illumination is flux per unit of area, we get from (11)

$$E=B\frac{df}{df'}\cos W\cdot d\Omega=B\cdot\frac{1}{M^2}\frac{S\cos^4 W}{e^2}. \tag{23}$$

A comparison of (14) and (22) gives for the case of an object at infinity

$$\frac{df}{df'}\cdot\frac{1}{e^2}=\frac{1}{F^2}. \tag{24}$$

Therefore

$$E=\frac{BS\cos^4 W}{F^2}\text{(object at infinity)}. \tag{25}$$

This expression holds for any size of the entrance pupil, since when integrating over the entrance pupil all the angles W are equal.

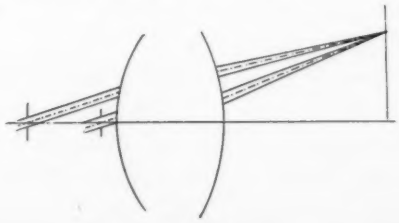


FIG. 6.

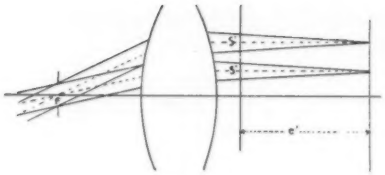


FIG. 7.

If we consider the pencil leaving the exit pupil we have

$$E = B\frac{df'}{df'}\cos W' d\Omega', \tag{26}$$

or by substituting for $d\Omega'$

$$E = \frac{BS'\cos^4 W'}{e'^2}. \tag{27}$$

DISCUSSION OF THE RESULTS

The equations (25) and (26) suggest the frequent superficial way of writing

$$E = E_0 \cos^4 W, \tag{28}$$

$$E = E_0 \cos^4 W'. \tag{29}$$

These equations obviously cannot both be right. A few examples, included in the paper by Slussareff,[4] show that (28) and (29) may be used only if proper precautions are observed.

Let us assume a lens system to be *fully* corrected for a parallel pencil and let us shift the diaphragm. Then for each position of the diaphragm the entering pencils are identical (Fig. 6). Since the lens is corrected for all these pencils the illumination must be equal for each of them. This is in agreement with Eq. (25) since B, S, and F are constant. But according to (29) the illumination should vary since W' is different for each position of the diaphragm. Equation (29) must therefore be wrong for this case. By using (27) and remembering that according to (22)

$$\frac{S'\cos^4 W'}{e'^2} = \frac{S\cos^4 W}{F^2} = \text{constant}, \tag{30}$$

we again get $E = \textit{constant}$ for all stop positions. If we wish to have an expression including the illumination at the center of the field for the case of a fixed diaphragm in front, then from formula (22):

$$S' = S\cdot\frac{e'^2}{F^2}\cdot\frac{\cos^4 W}{\cos^4 W'}, \tag{31}$$

and for the point on the axis ($W = W' = 0$),

$$S_0' = S\frac{e'^2}{F^2}. \tag{32}$$

Therefore

$$S' = S_0'\frac{\cos^4 W}{\cos^4 W'}. \tag{33}$$

We can look upon $(\cos^4 W/\cos^4 W')$ as being the distortion factor of the exit pupil. Using this value for S' we get from (27),

$$E = B\frac{S_0'}{e'^2}\frac{\cos^4 W}{\cos^4 W'}\cos^4 W', \tag{34}$$

for the point on the axis

$$E_0 = B\frac{S_0'}{e'^2}. \tag{35}$$

And finally

$$E = E_0\left(\frac{\cos^4 W}{\cos^4 W'}\right)\cdot\cos^4 W'. \tag{36}$$

THE SPECIAL CASE OF THE TELECENTRIC SYSTEM

If the diaphragm is at the front focus F (Fig. 7), then for all obliquities W the emerging principal ray will be parallel to the axis. According to (25) the illumination should decrease in proportion to $\cos^4 W$. On the other hand, according to (29) the illumination should be constant. This presents a paradoxical situation which can be resolved by the following argument.

Equation (22) gives for all obliquities:

$$\frac{S}{F^2}\cdot\cos^4 W = \frac{S'}{e'^2}\cos^4 W' = \frac{S'}{e'^2}, \tag{37}$$

where $d\Omega' = S'/e'^2$ is the solid angle corresponding to the exit pupil viewed from the respective image point.

[4] Reference 1. It was considered useful to give a more detailed discussion of these examples.

The exit pupil again varies with obliquity:

$$d\Omega' = \frac{S}{F^2} \cos^4 W, \tag{38}$$

thence the illumination is

$$E = \frac{BS'}{e'^2} = B \cdot d\Omega' = \frac{BS}{F^2} \cos^4 W. \tag{39}$$

Hence the conclusion from (29) is incorrect, and the illumination varies with the obliquity of the entering light.[1]

SYSTEMS HAVING AN INTERNAL DIAPHRAGM

In his paper Slussareff derives an expression for the illumination by a lens system with an internal diaphragm (Fig. 8). We shall briefly outline his procedure. In using (25) the unknown quantity is S since the diaphragm is inside. S is evidently the cross section in the object space of the parallel pencil which after refraction at the front half of the system just fills the diaphragm. The cross section shall be referred to a plane through the paraxial image A of A' produced by lens 1. This plane is of course in the object space. Because of the distortion of the entering rays, this cross section will not be a circle and its area, varying with obliquity, will differ from the area corresponding to the paraxial magnification at the points A and A'. Equation (22) can not be used because its derivation assumes a point to point imaging between object and image produced by the front part alone. Since the front part alone is not corrected this assumption is not allowable. The correct way of finding S would be to trace a number of skew rays which after refraction at lens *1* will go through the border line of the diaphragm opening. The intersection of those rays in the object space with the plane at A gives the area S. A contribution to this question has been made by an experimental investigation of Francis Washer.[5] The size of the entrance pupil was determined by illuminating the diaphragm from the rear, taking a photographic picture of the apparent entrance pupil, and then measuring the area with a planimeter. This was done for obliquities up to 45° half-angle by turning the lens around a vertical axis. The experiments with four lenses, of course, carried out for apertures where no vignetting occurred, gave the result that by turning the lens the size of the apparent entrance pupil generally does not simply change with the cosine of the obliquity. There were departures up to 20 percent. This change of the size of the entrance pupil is caused by the distortion of the rays at higher obliquities. The amount of the distortion depends upon the type of the lens and the position of the diaphragm.

In order to find an analytic expression for S, Slussareff uses Seidel's theory. Considering a point B of the diaphragm (generally not in the meridional plane) as the object point and taking a ray BL', we can use Seidel aberrations in order to derive the coordinates in the plane through A after refraction at lens *1* (light going from right to left). This can be done for all the points of the circumference of the diaphragm. The analytical expression thus derived permits to find by integration the area S. The assumptions tacitly made in this derivation are: (1) that all the rays of a parallel pencil coming from left to right unite after refracton at lens *1* in one point L' on the focal plane of lens *1* (stigmatic image), (2) that the image height CL' used in the quoted paper is $F \cdot \tan W$ (freedom from distortion of the front part alone), (3) that Seidel aberrations are sufficient for determining the coordinates of the rays in the plane through A.

Unfortunately none of these assumptions seems to be justified. It is extremely unlikely that the front part of a system will give a stigmatic image which is also free from distortion. It generally shows heavy aberrations which have to be compensated by the rest of the system. Furthermore, since the slope angles of the rays

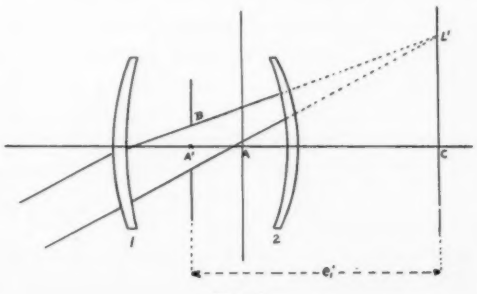


FIG. 8.

[5] F. Washer, Research Paper RP 1498, Nat. Bur. Stand.

at the diaphragm points are 30° or more (since we are interested in wide angle objectives) it is insufficient to use third-order theory for getting the coordinates at the plane through A.

Since all this effort is made for the purpose of deriving a correction to the paraxial value of the entrance pupil, it is evident that approximations introduced by the assumptions mentioned before are likely to lead to results which are much farther from the truth than the value which has to be corrected. Slussareff derives an expression which connects the size of the (variable) entrance pupil with the coefficient S_{II} in Seidel's theory. S_{II} is the coma coefficient for the front half of the objective, considering the diaphragm as the object. According to this formula it should be possible to increase the size of the entrance pupil by choosing suitable values for S_{II}. He therefore thinks it might be possible to design the system in such a way that the illumination is far better than for the known commercial systems. This optimistic conclusion seems not to be justified in view of the necessary assumptions.

Even if assumptions (1) and (2) were justified, assumption (3) alone (Seidel-theory) would make the whole derivation extremely doubtful. In this case the following consideration is of better use: According to (22) we have for the front half of the system

$$S = S'\left(\frac{\cos^4 W'}{\cos^4 W}\right)\frac{F_1^2}{e_1'^2}, \tag{40}$$

where W' is the ray slope in the diaphragm space, S' is the area of the diaphragm, and F_1 the focal length of the front part. The entrance pupil S is identical for the half-system and for the entire system. For the entire system we have:

$$E = \frac{BS\cos^4 W}{F^2}. \tag{41}$$

If we introduce for S its value given, we have

$$E = \frac{BS'F_1^2}{F^2 e_1'^2}\cdot\cos^4 W'. \tag{42}$$

In this case the illumination is proportionate to the fourth power of the cosine of the angle at the diaphragm. In order to improve the illumination the design has to be done in such a way that w' is reduced.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 4 APRIL, 1945

Illuminating and Viewing Conditions for Spectrophotometry and Colorimetry

ARTHUR C. HARDY
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received January 19, 1945)

ON June 17, 1942, the American Standards Association formally adopted, as the basic method of color specification, the procedure of computing color specifications from spectrophotometric data by means of the Standard Observer and Coordinate System recommended in 1931 by the International Commission on Illumination. The A.S.A. did not adopt Resolution 3 of the I.C.I., which recommended that, when measuring the color of reflecting materials, the incident light should make an angle of 45° with the normal and the direction of observation should be normal to the surface.[1] Since "Specifications of the spacial distributions of the incident and collected light are essential to the standardization of spectrophotometry," the A.S.A. did recommend that "until standard conditions are established by agreement, the particular conditions employed in each instance should be stated clearly." In view of some recent discussions in several quarters concerning the most desirable spacial distributions of the incident and collected light, it seems appropriate to review certain fundamental principles of optics that govern the choice of standard conditions.

CASE I—IDENTICAL SURFACE CHARACTERISTICS

Because of the fact that virtually every article of commerce is produced by methods of mass production, it might be assumed that the routine applications of the standard colorimetric procedure would involve samples having identical surface characteristics. In the textile field, for example, a single swatch of material may be the color guide for the production of millions of yards of cloth; and, ideally, swatches cut from any part of any bolt will be indistinguishable from the standard swatch, even under the severest scrutiny. A similar role is played by standard panels in the field of paints and industrial finishes. In short, the routine problems connected with mass production involve color comparisons between surfaces whose weave, texture, gloss, bronze, and other characteristics are nominally identical.

If two materials having identical surface characteristics are found to have identical spectral reflection characteristics under one mode of illumination and observation, they will be found to have identical spectral reflection characteristics under every mode of illumination and observation. Hence, in the absence of differences in surface characteristics, one mode is logically as adequate as another. Naturally, all modes are not equally advantageous in practice, but this aspect will be considered later.

CASE II—SLIGHTLY DISSIMILAR SURFACE CHARACTERISTICS

In the mass production of colored materials, variations from the standard surface characteristics frequently occur, owing to the operation of causes which are as diverse as the processes by which the materials are produced. These variations in surface characteristics are usually small, and they are not ordinarily caused by variations in the concentrations of the dyes or pigments employed. In other words, although relatively small variations in the dye or pigment content can produce significant color changes, they generally have little effect on the surface characteristics of the material. It follows that the standard colorimetric procedure should be, to the greatest extent possible, independent of differences in surface characteristics. Otherwise, the problem of those who use colorimetric procedures in the routine control of dye or pigment concentrations becomes enormously more complicated.

Of the various modes of illumination and observation that have been used, the one that is least sensitive to variations in surface charac-

[1] The text of this resolution is as follows: "Dans la mésure colorimétrique des matériaux réfléchissants, excepté dans le cas où des circonstances particulières indiquent qu'il en doit être autrement, la lumière incidente doit faire un angle de 45° avec le normale et la direction d'observation doit être normale à la surface de l'échantillon."

teristics employs substantially normal illumination and substantially diffuse observation. In a typical case, the sample is placed over a hole in an integrating sphere and is caused to be illuminated by a beam of light that enters through another hole on the opposite side of the sphere. The light reflected by the sample is integrated by the matte-white lining of the spheres; and, for the purpose of measurement, a small portion of the sphere-light is allowed to escape through a third hole. The amount of light that escapes is proportional, of course, to the amount of light reflected by the sample, provided all the holes in the sphere are small in relation to the size of the sphere.

That this normal-diffuse mode is insensitive to small variations in surface characteristics can be inferred from the fact that it leads to results which are only slightly affected, even by such extreme differences as are represented by curves *A* and *B* of Fig. 1. The sample in this case was an opaque slab of green Carrara glass, one face of which was first given a high polish. Half of this polished surface was then protected by a coating of paraffin while the other half was given a deep etch with hydrofluoric acid. After removal of the paraffin, the spectral reflectance of both areas was determined by means of a spectrophotometer[2] that employs the normal-diffuse mode. The resulting curves are nearly coincident, despite the fact that half of the surface was highly polished, while the other half was so matte[3] that no vestige of specularity was discernible.

As an example of the advantage resulting from the insensitivity of this mode to changes in surface characteristics, consider the problem of testing a material for weatherability. On exposure of the sample in a weather-ometer, the dyes or pigments which it contains may be faded, and/or the surface characteristics may be altered. However, these factors can be separated,

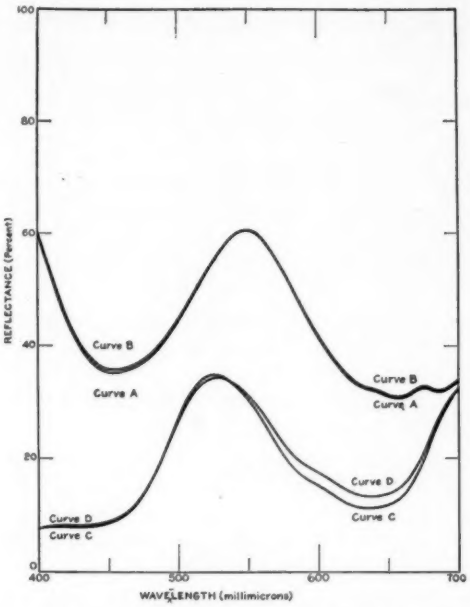


FIG. 1. Curves *A* and *B* indicate that, under the normal-diffuse mode, the spectral reflectance of a polished surface (curve *A*) is substantially the same as that of a matte surface (curve *B*) of the same material, in this case Carrara glass. Curves *C* and *D*, representing the spectral reflectance of a green plastic before exposure (curve *C*) and after exposure (curve *D*), indicate the usefulness of the normal-diffuse mode in analyzing the changes in color produced by weathering.

as curves *C* and *D* of Fig. 1 will illustrate. Curve *C* represents the spectral reflectance of a slab of green plastic before exposure, and curve *D* is the corresponding curve after exposure. The fact that the curves are nearly coincident in the spectral region from 400 to 530 millimicrons indicates that the yellow pigment, which is responsible for the absorption in this spectral region, was only slightly affected by the exposure. On the other hand, the disparity at wave-lengths longer than 530 millimicrons is a measure of the fading of the blue pigment, whose absorption maximum occurs at 635 millimicrons. These curves say nothing with respect to changes that may have taken place in the character of the surface; but that is properly a function of instruments for measuring gloss (or of a goniophotometer). The curves of Fig. 1 do indicate that, if the changes in surface characteristics are evaluated by means of a gloss-measuring instrument employing light in the spectral regions

[2] Arthur C. Hardy, J. Opt. Soc. Am. **25**, 305 (1935).

[3] The reflectance of a matte surface produced by grinding is usually higher than the reflectance of an etched surface, because of the debris that is forced into the interstices of the surface during the grinding. This debris consists of abrasive mixed with a large quantity of powdered glass, which is always white regardless of the original color of the glass. Thus the reflectance of a ground surface of black glass is decreased from more than 10 percent to approximately 4 percent by subsequent etching.

where curves *C* and *D* coincide, the indications of the gloss-measuring instrument will not be affected by the change in color of the sample. It need scarcely be remarked that, if the mode of illumination employed in the spectrophotometric measurements had been sensitive to changes in gloss, this sort of analysis could not have been made.

CASE III—LARGE DIFFERENCES IN SURFACE CHARACTERISTICS

Color matches are occasionally made between samples whose surface characteristics are nominally dissimilar. Such a situation frequently arises when an attempt is made to evaluate color by direct visual comparison with a set of "universal" color standards. As a concrete example, consider the use of the *Munsell Book of Color*, which contains a set of samples that are very nearly matte in their surface characteristics. Such samples, because of their matteness, can be compared with one another under almost any mode of illumination and observation without disturbing their apparent interrelationships. On the other hand, a glossy sample that matches one of the samples of the *Munsell Book of Color* under one mode will, in general, fail to match under any other mode. There is no escape from this inexorable principle; and every such set of color standards should be accompanied by a statement concerning the mode of illumination and observation to be used. In the case of the *Munsell Book of Color*, this mode has been tacitly specified by the fact that the Munsell System has recently been redefined[4] in terms of spectrophotometric data obtained through the use of the normal-diffuse mode. It goes without saying that the reciprocally related diffuse-normal mode is more readily adaptable to direct visual comparisons; and it is the equivalent of the normal-diffuse mode in every respect.

There is one special class of industrial problem that involves matching the color of materials whose surface characteristics are inherently different. As an example, a textile manufacturer may be asked to produce a dull-surfaced textile fabric that matches a shiny bone button in color. Obviously, the textile manufacturer can, in general, produce a match for only a single mode of illumination and observation. Indeed, if it were possible to produce a match under all conditions of illumination and observation, the button would always be invisible, except for its shadow, and would certainly fail of its decorative purpose.

These thoughts offer little comfort to the manufacturer who, although realizing that the match can be valid for only a single mode, must nevertheless produce a match that is valid under the mode which the customer had in mind but did not (or could not) specify. In 1929, the author described[5] a spectrophotometer in which the sample was illuminated normally and viewed at 45°; and the General Electric Company subsequently built a few instruments based on this geometry. Production of this model was discontinued because the experience of its users seemed to indicate that the mode of illumination and collection was not well suited to the general requirements of a color-measuring instrument. To cite a typical instance, there was objection[6] to the fact that a satin and a crepe which had been passed as a "commercial match" by the inspectors in a silk mill were found by this instrument to possess very different spectral reflection characteristics. It was surmised that the inspectors, when forced to compare two materials having such very different surface characteristics, inspect them under a variety of observing conditions and make some sort of mental average. On the assumption that this mental average corresponds to the result that an integrating sphere achieves in one operation, a viewing device was constructed in which the two materials could be compared visually under diffuse illumination and normal observation. When it had been demonstrated, after several months of trial, that the same inspectors would pass a satin and a crepe that appeared to match

[4] Final report of the O.S.A. Sub-Committee on the Spacing of the Munsell Colors, J. Opt. Soc. Am. **33**, 385 (1943).

[5] Arthur C. Hardy, J. Opt. Soc. Am. and Rev. Sci. Inst. **18**, 96 (1929).

[6] Of more serious consequence, when judged by present standards of accuracy, is the apparent variation in spectral reflectance produced by slight variations in the position of the sample. In the case of the 0°–45° instrument, merely removing a textile sample from the sample-holder and then replacing it for remeasurement usually changed the geometry significantly. This sensitivity to changes in geometry is at a maximum in the case of unidirectional light and vanishes when the light is diffuse.

under these conditions, the optical system of the spectrophotometer was redesigned; and the equivalent normal-diffuse mode has been embodied in all G. E. spectrophotometers constructed since 1933.

PRACTICAL CONSIDERATIONS

It is evident that, under a strict interpretation of the normal-diffuse mode, part of the light that is specularly reflected by a sample having a glossy or semi-glossy surface will escape through the hole by which the incident light enters the sphere. It is therefore good practice to design the integrating sphere in such a way that the incident light falls on the sample at a small angle (six degrees is a common value). This slight departure from strictly normal incidence is usually of no importance and it is advantageous in most instances to collect substantially all the light reflected by the sample. Indeed, if there is to be a standard mode of illumination and observation, the author believes that the specification should require "substantially normal illumination" of the sample and "substantially complete collection" of the light reflected by the sample. The term "substantially" recognizes the fact that, although the normal-diffuse conditions cannot be met exactly, they can be approximated so closely that, if an even closer approximation were made, the results would not be altered significantly.

The establishment of a standard mode of illumination and observation does not imply that no other mode should be used. It merely suggests that, when it has seemed desirable in special circumstances to use some other mode, a statement concerning the geometry of that mode should accompany the report of the data. As an example of such special circumstances, consider the problem of matching the color of confectioner's chocolate to one of its many standard shades. This matching is done by blending batches of different color until the molten mass has been brought to the desired shade. In evaluating the spectral reflectance of such a liquid sample, it is common practice to place the liquid in a glass cell, which reflects approximately 4 percent of the normally incident light at the air-glass interface and a fraction of 1 percent at the glass-chocolate interface. When the cell is placed against the sample-port of an integrating sphere arranged in the manner described above, this specularly reflected light illuminates a small area of the sphere-wall adjacent to the hole through which the incident light entered the sphere. If the sphere is provided with a port of slightly greater area, uncovering this port will allow the specularly reflected light to escape from the sphere. In other words, such an arrangement provides a means for eliminating surface reflections from the evaluation of the light diffusely reflected by the contents of the cell.

Spectrophotometry has come to be used so extensively as a basis for colorimetric specifications that it is easy to forget that there are many applications of spectrophotometry in which the color of the sample is of no interest whatever. Indeed, spectrophotometry has even been applied[7] to a study of the spectral absorption characteristics of the sensitizing dyes in an unexposed photographic emulsion, the color of which can obviously be of no interest to the photographer. To measure the spectral absorption characteristics of such an emulsion, the spectral reflectance and the spectral transmittance are measured for normal incidence and diffuse observation. If the collection of light is complete in both cases, the spectral absorptance is the difference, at each wave-length, between the sum of these quantities and unity. No other mode of illumination and observation can make this use of the principle of conservation of energy.

Since this section is intended to deal with practical considerations, it should be mentioned that the integrating sphere of a spectrophotometer is almost certain to be preceded in the optical train by a monochromator which places a definite limit on the amount of light-flux available. The best use of this flux is made by the normal-diffuse mode, since all the light reflected by the sample is collected within the integrating sphere. For obvious reasons, a mode of observation that collects only the flux reflected by the sample within a small solid angle is necessarily less efficient.

---

[7] J. A. Leermakers, J. Chem. Phys. 5, 889 (1937).

Design of a Printed Spectrum*

DAVID L. MACADAM
Kodak Research Laboratories, Eastman Kodak Company, Rochester, New York
(Received January 26, 1945)

A COLOR reproduction 19 inches long, of the solar spectrum, and a color plate representing the I.C.I. chromaticity diagram were published in the July 3, 1944 issue of *Life*. This paper is a brief account of the design and characteristics of those reproductions.

The plates used for printing the spectrum were prepared from positive copy on film base. These transparencies were made by exposing photographic films through the apertures of moving templates cut to the shapes shown in Fig. 1. The shapes of these templates were determined by preliminary tests which included preparation of plates that were used in printing separate and combined proofs of fourteen tones of each of the inks used in *Life*. The templates shown in Fig. 1 were moved at a uniform rate in the direction of the arrows, transversely to the wave-length scale, thus differently exposing various portions of the films. Development of these films produced transparent copy from which photoengravings could be made which would print the proper amounts of the colored inks throughout the spectrum.

In order to compensate for all complications of tone reproduction in the photographic steps as well as in the engraving process, calibration plates were engraved from a stepped-density transparency also exposed by the template procedure. The shape of the template used for calibration is shown in the lower portion of Fig. 2. The exposure and development were carefully standardized so that the density produced could always be reproduced by use of the same size of aperture. The densities of the areas exposed by the stepped template are shown above the corresponding apertures in Fig. 2. Each step of the calibration plates was 1 by 18 cm, the fourteen steps designated by letters in Fig. 2 being included in each 14- by 18-cm plate. Proofs were printed from these calibration plates, using each of the four inks with which the spectrum was ultimately printed. In addition to separate proofs, the inks were also printed together in pairs with the steps of the two impressions mutually perpendicular. Moiré patterns were eliminated by the use of calibration plates prepared with different screen angles. In this way proofs were obtained representative of the entire chromatic gamut obtainable with the inks used in printing *Life*.

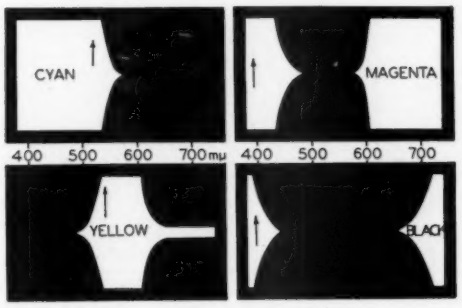


FIG. 1. Shapes of openings in templates used in making positive color-separation transparencies for *Life* spectrum.

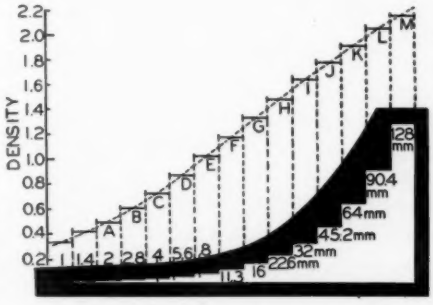


FIG. 2. Shape of opening in template used in making calibration stepped-density transparency for surveying colors obtainable with *Life* color printing process. Above: Densities of steps produced by corresponding apertures of template. Broken curve permits interpolation of aperture sizes necessary to produce intermediate densities.

The chromaticities produced when these proofs were illuminated with I.C.I. Standard Illuminant *C* are shown in Fig. 3. The letters along the boundaries correspond to the calibra-

* Communication No. 1021 from the Kodak Research Laboratories.

tion densities shown in Fig. 2 and designate the series of colors produced by the combination of the corresponding step of the particular ink indicated with all densities of the other two inks. Figure 3 indicates the entire gamut of chroma-

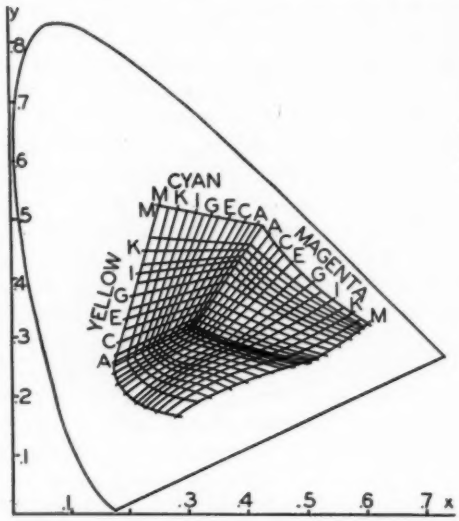


FIG. 3. Chromaticities obtained by superimposed proofs of pairs of *Life* printing inks are shown at intersections of curves in standard I.C.I. chromaticity diagram. Curves are designated by letters corresponding to those assigned in Fig. 2 to steps of calibration transparency.

ticities obtainable, each of which could be produced by a combination of not more than two of the printing inks. The densities of the transparencies which should be furnished to the engraver in order to produce any chromaticity (assuming that the procedure used in making the calibration engravings will be accurately repeated) could be determined by interpolation in Figs. 3 and 2. The sizes of the openings in the templates were interpolated from Figs. 3 and 2.

The spectrum was reproduced with the boundary chromaticities shown in Fig. 3. The transparencies submitted to the engraver had densities such that each portion of the printed spectrum should have had the dominant wave-length indicated by the printed wave-length scale. The openings in the templates, represented in Fig. 1, were determined from Figs. 3 and 2 to give the maximum purity of each dominant wave-length. For a check on processing, narrow strips along the edges of the films were masked and were subquently exposed with the stepped template shown in Fig. 2. Measurement of the processed stepped-densities proved that the photographic procedure had been accurately reproduced.

Although the printed spectrum was satisfactory, systematic errors appeared. These errors have been evaluated from one of the better copies of the spectrum, and were distinct from the obvious flaws of imperfectly printed copies. The errors in dominant wave-length of a good copy are shown in the upper portion of Fig. 4. These errors are shown also in terms of just noticeable wave-length differences in the lower portion of Fig. 4 in order to eliminate any exaggeration of the importance of the errors at the ends of the spectrum. These errors were attributed to failure to repeat accurately the procedure used in making the calibration engravings. The major differences between the calibration plates and the plates from which the spectrum was printed are indicated by the tone distortion curves shown in Fig. 5. The ordinates represent the densities of the original step-tablet transparency which gave the same colors in the calibration proofs as were obtained in portions of the published spectrum for which the submitted transparencies had the densities indicated by the abscissas. If the engraving process used in making the calibration plates had been repeated exactly, the curves in Fig. 5 would be coincident with the straight line inclined at 45 degrees to the horizontal, and the errors of dominant wave-length in the printed spectrum

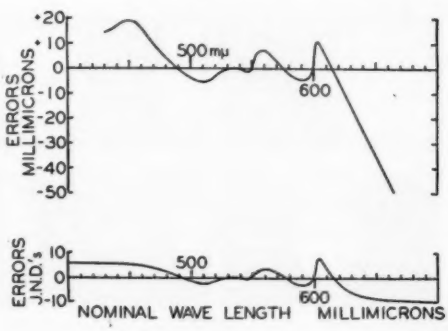


FIG. 4. Errors of dominant wave-length of a good copy of *Life* spectrum, shown as a function of nominal wave-length (according to printed scale). Below: Errors of dominant wave-length evaluated as multiples of just noticeable differences.

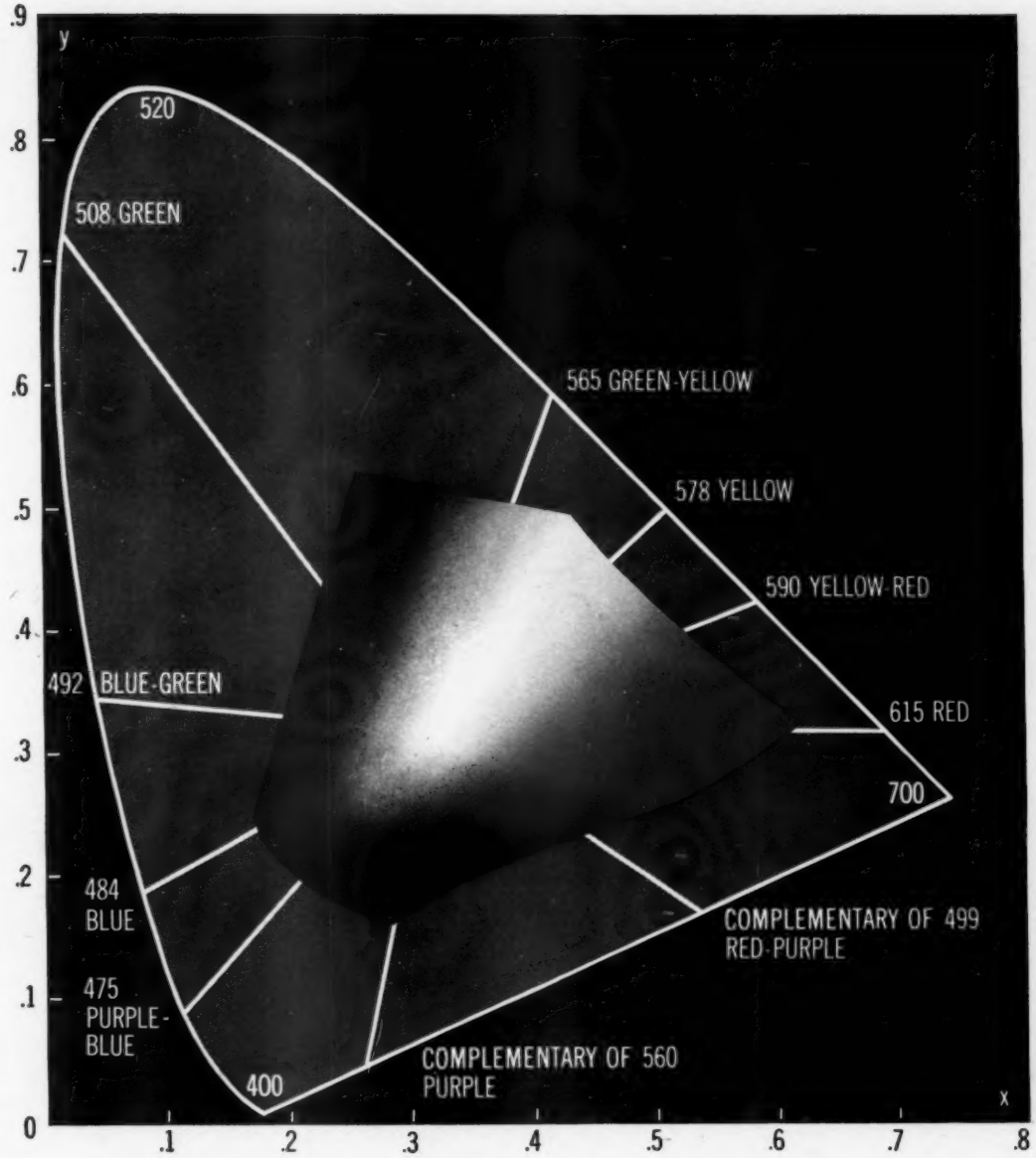


PLATE I. Distribution of colors in I.C.I. chromaticity diagram, showing gamut obtainable with a modern commercial process of color printing. The colors are approximately correctly distributed in the diagram when the plate is illuminated with daylight. (Inserts provided by *Life*, reprinted from issue of July 3, 1944.)

would be everywhere zero rather than as indicated by Fig. 4. The widely divergent curves shown for portions of the magenta and yellow plates are evidence of local treatment of the plates and account for the rather large errors at the ends of the spectrum.

No attempt was made to reproduce accurately the relative luminance of various wave-lengths of the solar spectrum. Adjustments for this purpose could have been made only by a sacrifice of purity or luminance of various portions of the spectrum. The most disappointing feature of the printed spectrum was the abruptness and unevenness of the black printer which was used to shade the ends of the spectrum to black. The decision to omit black shading from the major portion of the solar spectrum and from all of the line spectra was therefore fortunate.

Plate I, representing the distribution of chromaticities in the standard I.C.I. color-mixture diagram, was designed from Figs. 3 and 2. Positive transparencies having the required distributions of densities were prepared by pinhole photography of negative patterns built up of layers of 0.10 density neutral filter-films cut to the shapes indicated in Fig. 6. The manner of assembling the numerous pieces of neutral density film is shown in section in the upper left portion of Fig. 6. The assembled patterns were surrounded by opaque masks having the outlines indicated by dashed lines in Fig. 6, and were mounted on an illuminator. With a pinhole halfway between each pattern and the photographic film, and with a pinhole diameter one half the width of the widest step in the patterns, a photograph was obtained at unit magnification with all of the steps converted to smoothly shaded distributions. The pattern for the gray border was prepared directly from Fig. 3, and this, in turn, was used as an overlay to guide the engraver in cutting away the excess outer portions of the plates, which resulted from the diffused edges of the images. As in the case of the spectrum, there were systematic errors in the distribution of chromaticities in Plate I, caused by differences of tone reproduction between the calibration plates and the final printing plates. Plate I was designed to produce the proper chromaticities when illuminated with average daylight (I.C.I. Standard Illuminant *C*). No attempt was made to maintain equal luminous reflectance. The chromaticities were produced with not more than two inks at any one location in the diagram. This produced each chromaticity at the highest luminance possible with the inks employed. Errors owing to printing variations were held to a minimum by the use of the minimum possible number and amounts of inks Use of three or all four of the regular inks together could not increase the gamut of chromaticities, but would serve only to darken the colors.

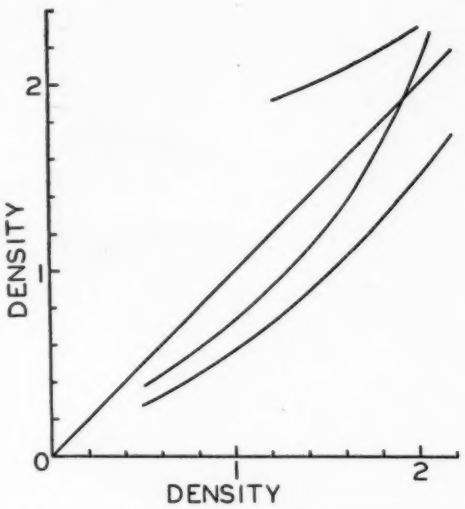


FIG. 5. Tone distortion curves of components of *Life* spectrum. Ordinate shows density of the step of the calibration transparency from which was obtained the same color as that printed in *Life* spectrum (at a wave-length for which the printer's copy had the density indicated by the abscissa). The straight line at 45° shows expected tone reproduction, which would have been obtained if the spectrum printing plates had been made exactly the same as the original calibration engravings. The curve which crosses the ideal line represents the distortions of all three printed colors between 536 mμ and 605 mμ. The upper curve shows characteristics of insufficiently etched portion of the yellow printing plate between 605 mμ and 700 mμ. These distortions are responsible for the orange appearance of long wave-length end of *Life* spectrum. The lowest curve shows characteristics of excessively etched portion of magenta printing plate between 400 mμ and 535 mμ.

ACKNOWLEDGMENT

I wish to ackowledge with gratitude the assistance of Herbert E. Goldberg, Walter Litten, Louis M. Condax, and Franklin C. Williams. I wish to express my appreciation of

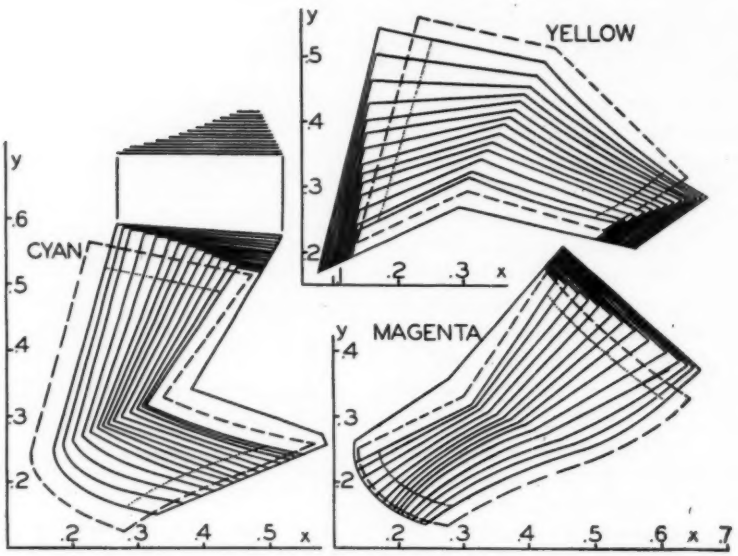


FIG. 6. Construction of patterns of superimposed pieces of 0.1 density neutral filter-film used in preparing Plate I. The shapes of the pieces were derived from Fig. 3, corresponding to the densities produced by photographing a trial assembly of 0.1 density neutral filter-films. Note the sectional view of the pile of filter-films shown at upper left. The coordinates shown for each pattern correspond to those in Fig. 3. Broken lines indicate the shapes of opaque masks which surrounded each pattern and concealed irregular ends of pieces of filter-film. These patterns were placed on an opal glass illuminated uniformly from the opposite side, and were photographed full size with a simple aperture (pinhole, no lens) having diameter equal to half the width of the widest step in the patterns. The engraved printing plates made from the resulting smoothly graded transparencies were trimmed corresponding to the dotted boundaries shown in the patterns.

the intelligent and helpful cooperation of the editors and printers of *Life* and of the courtesy and generosity of *Life* in making possible the insertion of Plate I.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 4 APRIL, 1945

Color of Soils

DOROTHY NICKERSON, *Office of Distribution, War Food Administration, Washington, D. C.*
K. L. KELLY, *American Pharmaceutical Association, Washington, D. C.*

AND

K. F. STULTZ, *The Engineer Board, Fort Belvoir, Virginia*

(Received January 2, 1945)

IN connection with the establishment of color standards and color names for soils[1] in the United States Department of Agriculture, a careful collection of several hundred samples of soil colors found in this country and its territories has been made. Mr. T. D. Rice and his associates spent many years in making sure that this collection covered all colors that should be included in the preparation of soil color names and standard charts. Because the color results are of interest for so many other purposes, they are reported here.

The color of 250 samples chosen to represent this color range was measured in containers such as are shown in Fig. 1, by visual observations of two observers[2] against the Munsell color charts by the mask method described by Judd and Kelly in 1939,[3] under 60 foot-candles of illumination supplied by Macbeth lamps at 7500°K. The average difference between readings by two observers, making independent judgments, and using hue charts 10 hue steps apart, was 0.6 ± 0.8 hue, 0.2 ± 0.2 value, and 0.3 ± 0.4 chroma steps. Average results in terms of Munsell book notations are shown by the small closed circles distributed on the 1YR to 7.5GY charts of Fig. 2.

In 1942 and 1943 representative samples taken from this series, together with samples selected by soil scientists to represent the color range of foreign soils, were measured wet and dry—in containers similar to those used in making visual judgments—on a G. E. recording spectrophotometer. Several samples were firmly packed so that measurements might be made on wet and dry samples both with and without glass covers, to provide data for correcting curves made under glass. A correction curve was adopted that checked theoretically very well for samples wet enough to allow no air space between glass and sample, and measurements of samples under glass were so corrected. Results of spectrophotometric data in the 400–700 mμ range for all samples measured are shown in Fig. 3. Typical soils, wet and dry, measured in the visual and infra-red, 400–900 mμ, are shown in Fig. 4. Munsell re-notations[4] for these samples are plotted in crosses and open circles (for wet and dry samples) on Fig. 2 both for domestic (small) and foreign samples (large). As may be seen, the color range for the domestic samples covers that of the foreign samples.

The average correction (based on 76 samples spectrophotometered) needed to change the Munsell notations obtained by visual comparisons to renotations (based on the same comparisons) was $0.2\pm0.3/0.2\pm0.2/0.3\pm0.1$. The average difference between renotations based on visual observations and renotations based on spectrophotometric methods for the same 76 samples was $1.2\pm1.0/0.2\pm0.3/0.6\pm0.2$.

FIG. 1. Container used for soil color measurements, shown with mask and charts used in making visual comparisons of color.

Detailed colorimetric data for these soil

[1] T. D. Rice, Dorothy Nickerson, A. M. O'Neal, and James Thorp, "Preliminary color standards and color names for soils," U. S. Department of Agriculture, Misc. Pub. No. 425 (September, 1941).

[2] Dorothy Nickerson and K. L. Kelly.

[3] D. B. Judd and K. L. Kelly, J. Research Nat. Bur. Stand. **23**, 355–385 (1939), illus. (RP 1239).

[4] S. M. Newhall, Dorothy Nickerson, and D. B. Judd, J. Opt. Soc. Am. **33**, 385–418 (1944).

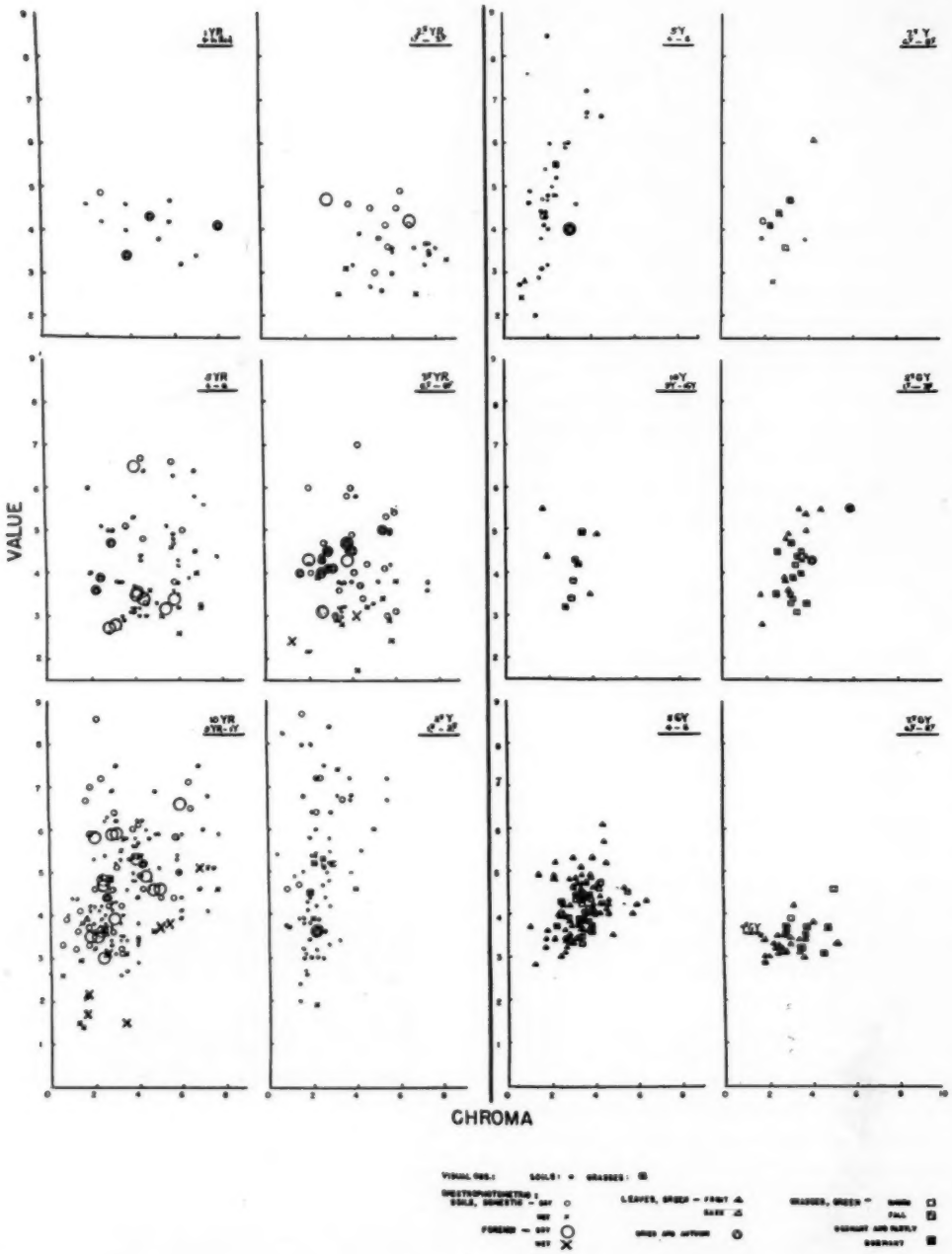


FIG. 2. Color distribution shown on a series of Munsell hue charts (1YR and redder; 2.5YR, etc. to 7.5GY) for visual observations and for calculations from spectrophotometric curves on a representative series of soil colors, with a few samples of foliage and grasses.

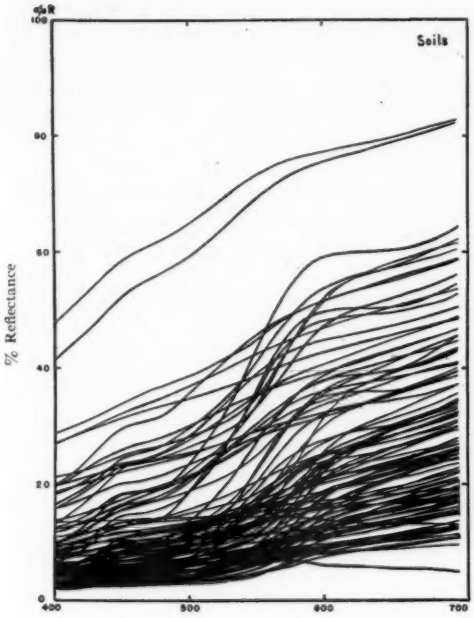


FIG. 3. Spectral reflectance curves of soil colors, 400 to 700 mμ.

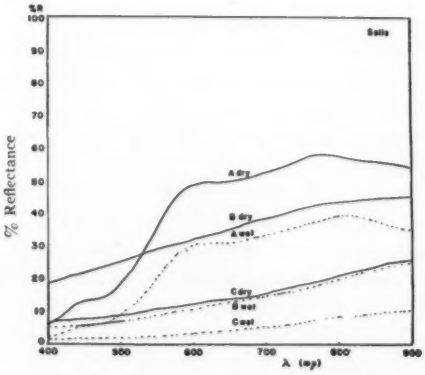


FIG. 4. Spectral reflectance curves of typical soils, wet (solid line) and dry (broken line), measured in the visual and infra-red regions, 400–900 mμ.

samples, with identifying name and geographical origin, have been compiled by the senior author and a few copies of the tables are available if needed.

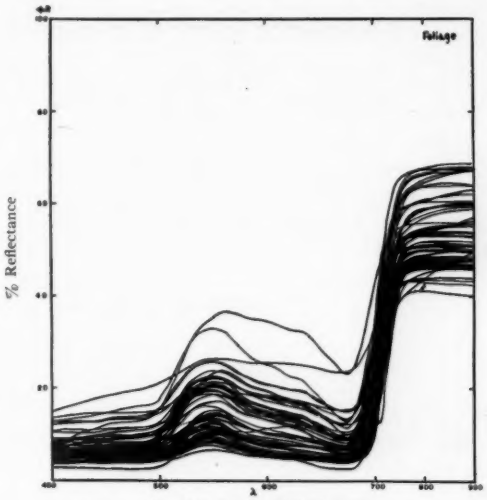


FIG. 5. Spectral reflectance curves of foliage samples, 400 to 900 mμ.

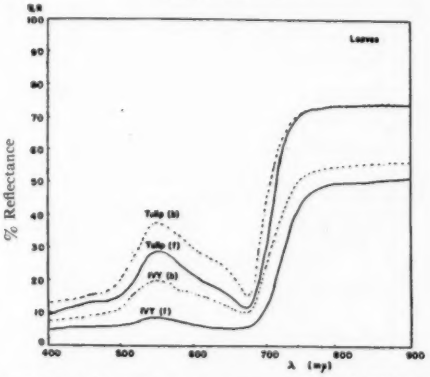


FIG. 6. Spectral reflectance curves of front (solid line) and back (broken line) of a very light and a very dark leaf.

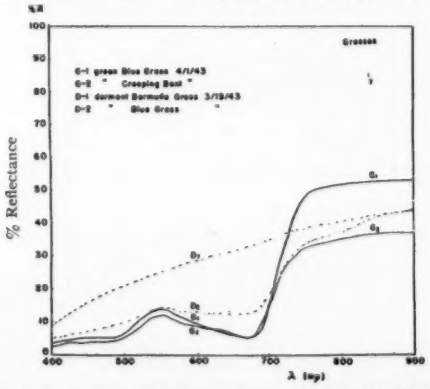


FIG. 7. Spectral reflectance curves of green (solid line) and dormant (broken line) grasses.

FOLIAGE AND GRASS COLOR

Renotations for a few samples of foliage and grass were obtained during the latter part of the soil color study, and because of their interest they are included in Fig. 2. Renotations for foliage colors were obtained from spectrophotometric curves such as are illustrated in Figs. 5 and 6. Renotations for grass samples were obtained from visual observation and from spectrophotometric curves such as are illustrated in Fig. 7.

Optical Society of America, Incorporated
Notice of Election of Officers

IN accordance with Article III, Section 6 of the By-Laws of the Optical Society of America, the following officers and directors are to be elected at or before the Annual Meeting in October, 1945.

President, Vice President — for a term of two years.
Secretary, Treasurer, 3 Members-at-large of the Board of Directors — for a term of four years.

The present officers are:

		Term expires at annual meeting in
President	—A. H. Pfund	1945
Vice President	—G. R. Harrison	1945
Secretary	—A. C. Hardy	1945
Treasurer	—W. W. Graeper	1945
Past President	—A. G. Worthing	1945
Editor	—G. R. Harrison	(Elected by Directors)

Directors-at-large:

Selig Hecht	1947
D. B. Judd	1945
D. L. MacAdam	1945
R. A. Sawyer	1945
Norman Wright	1947

In accordance with the By-Laws, Article III, Section 4:

> "No member at large of the board of directors shall be eligible for re-election."

D. B. Judd, D. L. MacAdam, and R. A. Sawyer are not eligible for re-election.

The following Committee on Nominations has been appointed and confirmed in accordance with Article III, Section 6 of the By-Laws:

A. G. Worthing, *Chairman*,
University of Pittsburgh,
Pittsburgh, Pennsylvania.

W. E. Forsythe,
Nela Park,
Cleveland, Ohio.

L. A. Jones,
59 Kodak Park,
Rochester 4, New York.

George S. Monk,
University of Chicago,
Chicago 37, Illinois.

W. B. Rayton,
Bausch & Lomb Optical Company,
Rochester 2, New York.

It will be the duty of this committee to make nominations for the above-mentioned election. The committee report is to be made not later than July 15, 1945.

Although, after the committee has made its report, an opportunity will be given for additional nominations by petition, members are urged, as soon as possible, to suggest names which should be considered by the committee. Suggestions may be sent to the Chairman or to any other member of the committee.

Only regular, honorary, and regular-emeritus members are eligible to election. A list of members was published in the July, 1944 issue of the *Journal of the Optical Society of America*. (Later transfers to regular membership will appear in the Secretary's Report for 1944, which is scheduled for publication in the June, 1945 issue of the Journal.)

ARTHUR C. HARDY, *Secretary*

Luther's Empirical Rule and the Quantum Theory of Exposure*

LUDWIK SILBERSTEIN
Eastman Kodak Company, Rochester, New York
(Received December 9, 1944)

THE property of the underexposed portion of the characteristic curves of photographic emulsions found empirically by Luther,[1] from measurements on forty different materials, and confirmed by Dr. L. A. Jones of these Laboratories, can be stated concisely as follows:

The slope g_i of the curve at the exposure corresponding to *the inertia* (i) is a *constant fraction* of gamma, that is, independent of γ and of the material, and approximately equal to *one half*.

It will be recalled that what, after Hurter and Driffield, is called the "inertia" of the material is the exposure $E=i$ corresponding to the intersection point of the Log E axis with the straight-line portion of the curve, more accurately, with the tangent at the inflection point of the curve, as explained in Fig. 1. Gamma is the slope of this tangent itself.

Our attention was recently drawn to Luther's result by Mr. Trivelli who has asked whether it could be derived from the equations of the characteristic curves given by our quantum theory of exposure. The purpose of the present paper is to show that the answer is, emphatically, in the affirmative. The Luther ratio g_i/γ for a manifold of quantic curves, to which corresponds a very wide range of γ-values, has turned out to be very nearly constant and not far from Luther's mark, 1:2.

Let $\bar{D}$ be the density at the inflection point of the curve, corresponding to the exposure $\bar{E}$. Then, by the definition of gamma and of i, Log $\bar{E}/i=\bar{D}/\gamma$ or, if y be a variable proportional to E, with y_i corresponding to $E=i$,

$$\text{Log}\,\frac{\bar{y}}{y_i}=\bar{D}/\gamma,$$

and therefore,

$$\text{Log}\,y_i=\text{Log}\,\bar{y}-\bar{D}/\gamma. \qquad (1)$$

This, of course, is quite general, independent of the form of the function $D=D(y)$. Again, whatever the form of this function determining the shape of the curve, the slope or gradient at any point of the D, Log E curve is

$$g(y)=\frac{dD}{d\,\text{Log}\,E}=\frac{E}{M}\frac{dD}{dE}=\frac{y}{M}D'(y),$$

where $M=\text{Log}\,e=0.4343$. Thus, in particular,

$$g_i=\frac{1}{M}y_iD'(y_i), \qquad (2)$$

where y_i is as in (1); and since

$$\gamma=(1/M)\bar{y}D'(\bar{y}),$$

the Luther ratio $L=g_i/\gamma$ becomes

$$L=\frac{y_i}{\bar{y}}\frac{D'(y_i)}{D'(\bar{y})}. \qquad (3)$$

If the exposure law, $D=D(y)$, be given, so is $D'(y)$; $\bar{y}$ is determined by $D'+\bar{y}D''=0$, whence also γ, $\bar{D}=D(\bar{y})$, and y_i by (1), so that (3) gives the numerical value of L.

Let us now evaluate the number L for characteristic curves or function forms $D(y)$ based on the quantum theory of exposure.

Any of these functions is of the form $D(y)=D_mF(y)$, where $F(\infty)=1$ and $D(\infty)=D_m$. By (3), the Luther ratio is manifestly independent

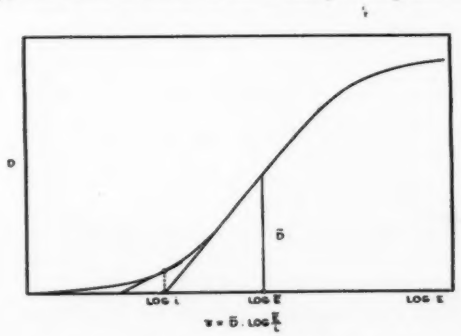


FIG. 1. Explaining the concept of inertia, $E=i$.

* Communication No. 1004 from the Kodak Research Laboratories.

[1] R. Luther, Trans. Faraday Soc. **19**, 340 (1923), as quoted by Dr. C. E. K. Mees in *The Theory of the Photographic Process* (The Macmillan Company, New York, 1942), pp. 209–210.

of D_m. Hence, without loss to generality, we may assume a unit limiting density, $D_m=1$.

In order to cover a wide class of emulsions, let us consider the two extreme cases, a *uniform* emulsion, consisting of equal grains,* and an "*exponential*" emulsion whose size-frequency distribution, that is, is exponential, $f(a)=(N/\bar{a})e^{-a/\bar{a}}$ (N=number of grains over unit area, $f(a)da$ =number of grains of size a to $a+da$). The former case will be approached by emulsions with a narrow range of grain sizes, and the latter by emulsions with a very broad size-frequency curve.

I. UNIFORM EMULSIONS

In our usual notation, $y=na\epsilon$, where n is the number of quanta incident, per unit area, and ϵ is the "intrinsic sensitivity" of a grain, a numerical fraction of the order 10^{-4} or so. If r be the number of quanta required for the developability of a grain, then

$$D(y)=1-e^{-y}\left[1+y+\frac{y^2}{2!}+\cdots+\frac{y^{r-1}}{(r-1)!}\right], \quad (4)$$

and, therefore,

$$D'(y)=\frac{y^{r-1}e^{-y}}{(r-1)!}. \quad (4')$$

The inflection point is determined by $D'+yD''=0$ or $ry^{r-1}-y^r=0$, which gives $\bar{y}=r$ (as in earlier papers). Thus,

$$\gamma=\frac{(r/e)^r}{M(r-1)!}, \quad (4'')$$

and our general formulae, (1) and (3), become

$$\text{Log}\,\frac{y_i}{r}=-MD(r)(r-1)!\left(\frac{e}{r}\right)^r, \quad (1a)$$

and

$$L=\left(\frac{y_i}{r}\right)^r\cdot\exp\,(r-y_i),$$

or, even by (1a),

$$\text{Log}\,L=M\left\{r-y_i-r!D(r)\left(\frac{e}{r}\right)^r\right\}, \quad (3a)$$

where $D(r)$ is directly given by (4). These formulae hold for any quantum number, $r=1, 2, 3$, and so on.

Start with the simplest case, $r=1$. Then, by (1a), recalling that 0! stands for 1, and $D(1)=1-e^{-1}$,

$$\text{Log}\,y_i=-M(e-1)=-0.746=\bar{1}.254,$$

to three decimals. Whence $y_i=0.179$ and, by (3a),

$$\text{Log}\,L=M(1.821-e)=-0.390=\bar{1}.610.$$

Thus, the Luther ratio for our *one*-quantic curve becomes

$$L_1=0.407.$$

This looks encouraging, not being much less than his "one-half." The corresponding gamma is, by (4''),

$$\gamma_1=\frac{1}{Me}=0.847.$$

The main point, however, consists in seeing whether L remains nearly constant when we pass to higher values of r and thus to greater γ.

Consider, then, the case of $r=2$. By (1a) and (4),

$$\text{Log}\,\frac{y_i}{2}=-\frac{M}{4}D(2)e^2=-\frac{M}{4}(e^2-3)=\bar{1}.523,$$

whence $y_i=0.668$ and, by (3a),

$$\text{Log}\,L_2=M(2.832-\tfrac{1}{2}e^2)=-0.863M$$
$$=-0.375=\bar{1}.625.$$

Thus, for a *two*-quantic curve,

$$L_2=0.422,$$

but slightly greater than L_1, although

$$\gamma_2=\frac{4}{Me^2}=\frac{4}{e}\gamma_1$$

is 1.47 times the previous γ.

In much the same way we might compute the Luther ratio for $r=3$, 4, and so on, only to find that L_3-L_2 is yet smaller than $L_2-L_1=0.015$, and so on. Instead of this laborious climbing, let us jump at once to enormous quantum numbers r, nay, to $r\to\infty$. As explained in an earlier paper,[2]

* Each of size (projected area) a.

[2] L. Silberstein, J. Opt. Soc. Am. **32**, 474 (1942).

the function (4) then becomes

$$D(y)=\tfrac{1}{2}\left[1+\Phi\left(\frac{y-r}{(2y)^{\frac{1}{2}}}\right)\right], \qquad (4\infty)$$

where Φ is the error-function. Thus, $\bar{y}=r$, as before,

$$\bar{D}=D(r)=\tfrac{1}{2}, \quad \gamma=\frac{1}{M}\left(\frac{r}{2\pi}\right)^{\frac{1}{2}},$$

and Eq. (1) becomes

$$\text{Log } y_i=\text{Log } r-M\left(\frac{\pi}{2r}\right)^{\frac{1}{2}}, \qquad (A)$$

and Eq. (3a), which holds for any r,

$$\text{Log } L=M\left\{r-y_i-\tfrac{1}{2}r!\left(\frac{e}{r}\right)^r\right\}.$$

Since r is an enormous number, $r!(e/r)^r=(2\pi r)^{\frac{1}{2}}$, so that we now have

$$\text{Log } L=M\left\{r-y_i-\left(\frac{\pi r}{2}\right)^{\frac{1}{2}}\right\}. \qquad (B)$$

Equations (A) and (B) hold, nearly enough, even for moderately great quantum numbers, such as $r=10$ or 20, for such is the case of the Stirling approximation implied throughout. The elimination, however, of y_i between (A) and (B) when $r\to\infty$ calls for some care. By (A), we have, passing from Briggsian to natural logarithms,

$$\log\frac{y_i}{r}=-\left(\frac{\pi}{2r}\right)^{\frac{1}{2}}, \quad y_i=r\exp-\left(\frac{\pi}{2r}\right)^{\frac{1}{2}}.$$

Hence, for great r,

$$y_i=r\left(1-\left(\frac{\pi}{2r}\right)^{\frac{1}{2}}+\frac{\pi}{4r}+\text{higher terms}\right),$$

so that

$$r-y_i=\left(\frac{\pi r}{2}\right)^{\frac{1}{2}}-\frac{\pi}{4},$$

and Eq. (B) becomes

$$\text{Log } L_\infty=-\frac{\pi M}{4},$$

or, more elegantly,

$$L_\infty=\exp-\frac{\pi}{4}.$$

Numerically, to four figures,

$$L_\infty=0.4560,$$

while γ is infinite, of the order of $\sqrt{r}$. To sum up, we have $L_1=0.407$, $L_2=0.422$, $\cdots L_\infty=0.456$, $\gamma_1=0.85$, $\gamma_2=1.25$, $\cdots\gamma_\infty=\infty$. In fine, the ratio L, whose measurements are scarcely reliable to more than two decimals, increases only from 0.41 to 0.46 when, with increasing quantum number, gamma rises from 0.85 beyond every limit.

Such is the property of quantic curves for a uniform or quasiuniform emulsion. Let us now pass to the other extreme mentioned in the preamble.

II. EXPONENTIAL EMULSIONS

As found before,[3] the r-quantic curve for such an emulsion is expressed by

$$D(y)=\frac{y^r(r+1+y)}{(1+y)^{r+1}}, \qquad (5)$$

where $y=n\bar{a}\epsilon$. The derivative of this function is

$$D'(y)=r(r+1)\frac{y^{r-1}}{(1+y)^{r+2}}, \qquad (5')$$

and the inflection point,

$$\bar{y}=\frac{r}{2}, \quad \bar{D}=\frac{(3r+2)r^r}{(r+2)^{r+1}}.$$

These results hold for any integer r (1, 2, 3, etc.) as the number of quanta required for the developability of a grain.

The value of gamma which follows directly from (5′) is

$$\gamma=\frac{1}{M}\cdot\frac{4(r+1)r^{r+1}}{(r+2)^{r+2}}. \qquad (5'')$$

Thus, Formula (1) for the inertia-exposure becomes

$$\text{Log } y_i=\text{Log}\frac{r}{2}-\frac{M}{4}\cdot\frac{(r+2)(3r+2)}{r(r+1)}, \qquad (1b)$$

[3] L. Silberstein, J. Opt. Soc. Am. **28**, 456 (1938).

and the Luther ratio, (3), assumes the form

$$L=\frac{(r+2)^{r+2}}{4r^r}\cdot\frac{y_i^r}{(1+y_i)^{r+2}}. \qquad (3b)$$

For $r=1$, we have

$$\text{Log } y_i=-\text{Log } 2-\frac{15}{8}M=\bar{2}.8867, \quad y_i=0.0770,$$

and

$$L_1=\frac{27}{4}\frac{y_i}{(1+y_i)^2},$$

or, to three decimals,

$$L_1=0.416,$$

remarkably close to the one-quantic ratio (0.407) found above for a uniform emulsion. The corresponding gamma is now

$$\gamma_1=\frac{8}{27M}=0.682.$$

Similarly, we find, by (1b) and (3b), for $r=2, 3, 4, 5$, respectively, $y_i=0.264, 0.476, 0.700, 0.927$ and $L_2=0.435$, $L_3=0.447$, $L_4=0.453$, and $L_5=0.457$. As before, the Luther ratio increases very slightly, and at a decreasing rate, with increasing quantum number, although γ grows conspicuously, *viz.*, $\gamma_2=3/8M=0.863$, $\gamma_3=0.955$, and so on.

Actually, the experimental curves of most emulsions call only for $r=1$ to 3. For the sake of completeness, however, let us compute the Luther ratio for very great quantum numbers, as in the case of uniform emulsions.

For $r\to\infty$, Formula (1b) tends to

$$\text{Log } y_i=\text{Log }\frac{r}{2}-\frac{3}{4}M,$$

or

$$y_i=\frac{r}{2}e^{-\frac{3}{4}}=\bar{y}e^{-\frac{3}{4}},$$

an interesting relation. Formula (3b) may be rewritten

$$L=\frac{1}{4}\frac{r^2}{y_i^2}\frac{\left(1+\frac{2}{r}\right)^{r+2}}{\left(1+\frac{1}{y_i}\right)^{r+2}},$$

and thus, for $r\to\infty$,

$$L=\frac{e^2}{4}\cdot\left(\frac{r}{y_i}\right)^2 e^{-r/y_i}.$$

But we have just found that $r/y_i=2e^{\frac{3}{4}}$. Therefore,

$$L_\infty=\exp\left[\frac{7}{2}-2e^{\frac{3}{4}}\right],$$

or

$$\text{Log } L_\infty=\left(\frac{7}{2}-2e^{\frac{3}{4}}\right)M=-0.734M,$$

to three decimals. This amounts to $\bar{1}.6812$, so that the limit of the required ratio for an ever growing quantum number is

$$L_\infty=0.480,$$

not much greater than L_5, and remarkably close to one half. The corresponding value of gamma is, by (5″),

$$\gamma_\infty=\frac{4}{Me^2}=1.247.$$

Thus, for the one-quantic, two-quantic, and higher curves of an "exponential" emulsion, we have the set of values:

$r=1$	2	3	4	5 ⋯	∞
$L=0.416$	0.435	0.447	0.453	0.456	0.480
$\gamma=0.68$	0.86	0.955	1.035	1.093	1.25 .

While gamma is almost doubled, the Luther ratio increases only from 0.42 to 0.48. Its values, moreover, are much the same as for a uniform emulsion.

Putting both sets of results together, we may say that all values of the ratio $L=g_i/\gamma$ are contained within *narrow* limits,

$$0.41\leq L\leq 0.48,$$

although the corresponding γ covers an infinite range, from 0.68 onwards. The ratio L, moreover, is not much less than one half. In fine, the empirical rule found by Luther holds with remarkable accuracy for all quantic curves constructed theoretically for uniform and for "exponential" emulsions. Since these represent the extreme cases of a very slender and a very broad

size-frequency curve,* our result can be expected to hold as well (and even more closely) for all intermediate emulsions, with a small or moderate ratio of standard deviation to average grain size.

We have, so far, considered the characteristic curves, of both types, expressed by a *single* quantic term (with any value of the quantum number r). As has been shown in a recent paper,[4] the actual characteristic curves of a good many emulsions are well represented by such single terms, namely, with an exponential size distribution. The great majority of emulsions, however, call for a *superposition of two* such terms, with $r=1$, 2, or 3; in most cases, both two-quantic, i.e.,

$$D(y)=\alpha F(y)+\beta F\left(\frac{y}{m}\right), \tag{6}$$

where

$$F(y)=\frac{y^2(3+y)}{(1+y)^3}$$

and Log m is the separation of the centers of the two component curves. It may be desirable, therefore, to determine the Luther ratio for such a compound curve. Its inflection point (or points), however, is (are) determined by an equation of the sixth degree,[4] which for $\alpha\neq\beta$, is not soluble in a round form. Only in the case of equipartition, i.e., for $\alpha=\beta$, does that sextic have the simple root $\bar{y}=\sqrt{m}$ (as well as the physically useless one, $y=-\sqrt{m}$), and the remaining quantic splits into two quadratic equations, which give an additional pair of inflection points placed symmetrically about the central $\bar{y}=\sqrt{m}$, if m is great enough. For smaller separations of the components, $\bar{y}=\sqrt{m}$ is the only real inflection point of the resultant curve.

Graphically, of course, the Luther ratio can easily be found for any curve of the type (6). (This may be left to the reader.) For an analytical determination of this ratio, we must be content with the special case of equipartition, $\alpha=\beta$, and since (as above) $D_m=1$, $\alpha=\beta=\frac{1}{2}$. Thus the equation of our curve becomes

$$D(y)=\frac{1}{2}\left[F(y)+F\left(\frac{y}{m}\right)\right], \quad F(y)=\frac{y^2(3+y)}{(1+y)^3},$$

where the constant parameter m may have any value, say $m>1$. The curve (D, Log y) is rigorously symmetrical about its inflection point $\bar{y}=\sqrt{m}$, $\bar{D}=\frac{1}{2}$.*

Formula (1) becomes

$$\text{Log } y_i=\tfrac{1}{2}\text{ Log } m-\frac{1}{2\gamma}. \tag{7}$$

It remains to determine γ and the slope $g_i=g(y_i)$. By (5′), with $r=2$,

$$D'(y)=\frac{1}{2}\left[F'(y)+\frac{1}{m}F'\frac{(y)}{m}\right]$$

$$=3y\left[\frac{1}{(1+y)^4}+\frac{m^2}{(m+y)^4}\right].$$

Hence, the slope at any point of the curve,

$$g(y)=\frac{3y^2}{M}\left[\frac{1}{(1+y)^4}+\frac{m^2}{(m+y)^4}\right], \tag{8}$$

and in particular, for $\bar{y}=\sqrt{m}$,

$$\gamma=\frac{6m}{M(1+\sqrt{m})^4}. \tag{9}$$

This gives, for $m=1$, when the two components coincide and their resultant is a simple two-quantic curve, $\gamma=\frac{3}{8}M$, as before. The corresponding Luther ratio, as found above, is $L=0.435$.

Now, to see the effect of a considerable separation of the component curves, take $m=4$. Then, $\bar{y}=2$ and, by (9) and (7),

$$\gamma=0.682, \quad y_i=0.370.$$

Thus, by (8), $g_i=0.309$ and, the required ratio,

$$L=0.454,$$

but slightly greater than for $m=1$, although γ is now considerably smaller. In order to reduce γ still more, let us take $m=9$. Applying again

* Standard deviation (σ) a negligible fraction of mean grain size ($\bar{a}$), on the one hand, and $\sigma=\bar{a}$ on the other hand.

[4] L. Silberstein and A. P. H. Trivelli, J. Opt. Soc. Am. **35**, 93 (1945).

* In fact, $D(\sqrt{m})=\frac{1}{2}[F(\sqrt{m})+F(1/\sqrt{m})]=\frac{1}{2}$, since the function F has the property $F(y)+F(1/y)=1$.

Formulae (9), (7), and (8) and gathering the results, we now have the set of corresponding values:

$m=1$	4	9
$L=0.435$	0.454	0.428
$\gamma=0.86$	0.68	0.49 .

The variations of L are, to all purposes, negligible, in spite of a considerable drop of gamma. This example of a compound characteristic curve will suffice to illustrate our point, i.e., the pronounced stability of the theoretical ratio $L=g_i:\gamma$.

The experimental validity of Luther's rule is, undoubtedly, well established. The fact, therefore, that this rule or property follows rigidly from the theoretical equations of quantic curves of various types, and over a very wide range of γ-values, may be fairly claimed as one more piece of evidence for the quantum theory of photographic exposure.

Radiating Characteristics of Tungsten and Tungsten Lamps. A Correction

[J. O. S. A. 35, 108 (1945)]

W. E. Forsythe and E. Q. Adams
General Electric Company, Nela Park, Cleveland, Ohio

THE data given in Table III of this article which appeared in the February issue of the Journal of the Optical Society of America are on a basis of an intensity of 1000 candles rather than 100 candles as stated. This error occurs in the heading of Table III and also occurs in the first paragraph on page 111 and in the last three paragraphs on page 112.

A Photographic Lens System

A. ESTELLE GLANCY
American Optical Company, Southbridge, Massachusetts
(Received January 3, 1945)

IN the Nov. 1940 issue of the Journal of the Optical Society of America the writer published a description of a telescope objective characterized by a wide aperture ratio which was attained by the use of a plano meniscus lens element preceding the system.

Following is a description of a photographic lens system incorporating a very weak meniscus for the same purpose under different conditions.

As a basic system on which to build improvement of the spherical aberration a modified type of Cooke system was adopted, largely because some previous calculations by the writer were already available and also because this system has two air spaces either of which is sufficiently long for the interposition of an additional lens.

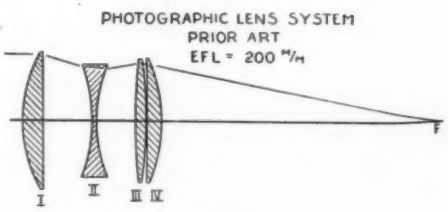


FIG. 1.

TABLE I. Photographic lens system. Prior art.

I.	+85.180 mm +1317.5	12.03 mm 30.02	
II.	−109.71 +84.946	2.47 24.06	$EFL_D = 200.00$ mm $BF_D = 165.92$
III.	+332.29 −332.29	6.07 0.20	
IV.	Plano −95.638	9.56	
I.-II.-IV.		Pittsburgh *DBC*−2 *C* 1.60774 *D* 1.61088 *F* 1.61842 *G* 1.62453	$\nu_D = 57.2$
II.		Pittsburgh *EDF*−1 *C* 1.64359 *D* 1.64900 *F* 1.66279 *G* 1.67458	$\nu_D = 33.8$

This system is given in U. S. Patent 1,739,512, (Lee, 1929). In place of the glasses specified, American glasses which are very similar are substituted. Also an error in the patent was corrected. The second surface is concave, not convex, as indicated in patent.

The significance of the new system lies in the fact that the zonal variation of the spherical aberration is removed to such an extent that the new system will give better images on the plate, as regards spherical aberration and sine condition, than the Lee system, even though the focal length of the improved system is twice as great as the Lee system. For this reason the data are given for 200 mm focal length rather than the usual 100 mm length.

The following figures (Fig. 1, Fig. 2) and Table I show essentially the Lee system.

It will be observed that the zonal variation of the spherical aberration and the sine condition are too great when the focal length is as much as 200 mm.

The improved system (as illustrated in Figs. 3, 4, and Table II) follows:

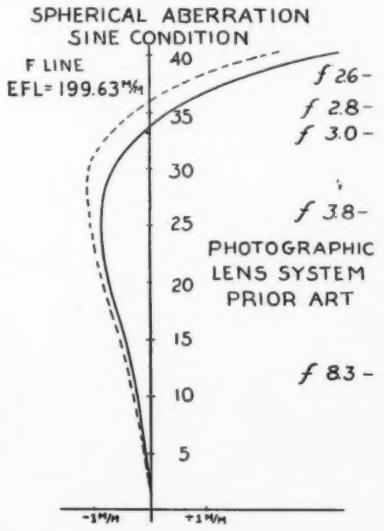


FIG. 2.

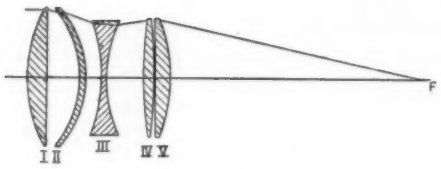


FIG. 3.

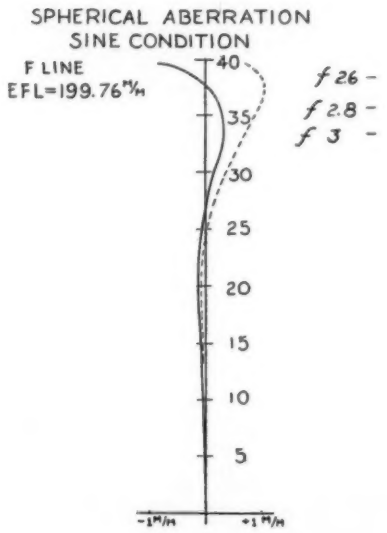


FIG. 4.

Although the meniscus element has only 3 percent of the power of the whole system it changes the axial color and the first-order spherical aberration. To bring these back without changing the field curvatures, primary and secondary, the system was corrected differentially. For this reason the surface curvatures and separations in the two systems are modified.

It is not certain whether the aperture ratio could be further extended. There are indications both pro and con. However, the amount of calculation which would be necessary to prove this point would not be justified.

In addition to the graphs shown here, extended calculations prove that the great improvement in spherical aberration and sine condition are attained at no expense to color and field curvatures and chromatic variation of the spherical aberration.

TABLE II.

I.	+88.981 mm	11.76	
	Plano	19.56	
II.	−63.727	3.31	
	−62.848	10.11	
III.	−105.27	3.30	
	+81.507	22.63	
IV.	+327.01	5.93	$EFL_D = 200.00$ mm
	−429.28	0.20	$BF_D = 165.92$
V.	+801.84	9.34	
	−97.538		
I.-IV.-V.		Pittsburgh *DBC*−2	
		C 1.60774	
		D 1.61088	$\nu_D = 57.2$
		F 1.61842	
		G 1.62453	
II.		Pittsburgh *C*−1	
		C 1.5203	
		D 1.5230	$\nu_D = 58.8$
		F 1.5292	
		G 1.5343	
III.		Pittsburgh *EDF*−1	
		C 1.64359	
		D 1.64900	$\nu_D = 33.8$
		F 1.66279	
		G 1.67458	

Election of Honorary Member of the Optical Society of America

THE Board of Directors of the Optical Society of America, at its meeting held in New York City on October 19, 1944, elected to honorary membership in this Society Professor R. W. Wood of Johns Hopkins University.

Robert Williams Wood came into the field of optics, where he is an unquestioned world leader, by the somewhat devious route of a preparatory school devoted to classical instruction, and entrance to Harvard University by examinations in which the required Greek and Latin were failed, but the unusual brilliancy of his showing in science, which he had worked up by himself, carried him in. Even then he did not at once find himself, for his first academic work in science was in chemistry, and it was only after several years in chemical study at Chicago and John Hopkins that he turned definitely to physics. It was while holding the position of instructor in physics at the University of Wisconsin that he broke into optics with his ingenious method of "diffraction color photography" in which three differently spaced diffraction gratings were utilized in place of color filters to produce red, green, and blue colored images. It was perhaps this work which suggested to Johns Hopkins University—the home of the famous Rowland grating—that he be called to that institution, as successor to Rowland, in 1904. There he has remained ever since, but not statically, for frequent journeys to Europe for periods of work in foreign laboratories have been evidenced by numerous papers in collaboration with physicists in France and Germany, by lectures at the Royal Institution and before the British Association. The esteem in which he is held abroad is shown by his foreign membership in the Royal Society of London, and in the Academies of Italy and Russia.

At Johns Hopkins, Wood's greatest work has undoubtedly been his extensive pioneer studies of the fluorescence of gases and vapors. In this work numerous new phenomena were discovered which have constituted the working data for all

Portrait of Robert W. Wood, taken at the time he was called to Johns Hopkins University.

c/o Harris & Ewing

Professor R. W. Wood, elected honorary member of the Optical Society of America.

modern theoretical work on optical spectra. Resonance radiation, the quenching of fluorescence by admixture of other gases, the effects of magnetic fields, may be mentioned as a few among the many significant phenomenon discovered or investigated by Wood. In these researches he explored the spectrum from the ultraviolet to the infra-red, enriching our knowledge by novel and ingenious techniques. Representative of his skill as an experimenter may be cited his disclosure of fifty-eight lines in the principal series of sodium in place of the much smaller number which previous experimenters had been able to resolve. These researches are described in numerous papers in the chief physical journals, noteworthy not only for their content, but for the brilliant and entertaining literary presentation. They may also be read in Wood's *Physical Optics*, a book which has appeared in three progressively larger editions, and is an indispensible fixture of every laboratory of optics.

Wood's early interest in the diffraction grating has persisted during his whole period of application to physical optics, and to the development and use of the grating he has made important contributions. While "freak" gratings, which threw light more into one order of diffraction than another were previously known, Wood has brought what was before an occasional accident under rigid control, and by purposeful shaping of the ruling tool and the grating surface, produces gratings in which practically all the light goes into one order. Carrying this technique into gratings of wide spacing he developed the "echelette" grating, a powerful tool for study of the infra-red. To other optical contributions in the field of astronomy he has recently added the development of large size built-up grating replicas suitable for use with the world's greatest telescopes. For this work, done since his official "retirement," he lately received the Draper medal of the National Academy of Sciences.

Of the many honors and awards which Professor Wood has received, mention may be made, as of special interest in connection with his optical work, of the Rumford medals both of the Royal Society of London, and of the American Academy of Arts and Sciences of Boston, and the Frederic Ives Medal of the Optical Society of America.

HERBERT E. IVES